A Knowledge Management Architecture for JPL

Knowledge Management Study Team

James Doane, Leader
Susan Hess
Lynne Cooper
Jeanne Holm
Donald Fuhrman
James U'Ren

January 15, 1999

Jet Propulsion Laboratory
California Institute of Technology

JPL Publication 99-18

The architecture in this document represents the work of many people. The core team, noted on the cover, coordinated, condensed, and analyzed the valuable comments and suggestions from across the Lab, as well as from outside consultants. The team would like to particularly acknowledge and thank those people who took extra efforts in contributing to this study.

- Pat Beauchamp—Center of Excellence, Project Management
- Larry Bowen—ISO 9000, Configuration Management
- John Casani—DMIE, ISO
- Ed Chow—Technologist
- Kathy Dumas—Science
- Susan Foster—Electronic Publications
- Peg Frerking—Program and projects
- Michael Gunson—Lead Scientist
- Kathleen Hardcastle—Human Resources
- Michael Hooks—Archives
- Steve Hughes—Science Data Products
- Steve Jenkins—EIS
- Margaret Johnson—Project Planning
- Pat Kleinhammer—Networks/Security
- Michael Kolar, Dan Crichton, Jason Hyon, Julie Reiz, Kris Blom, Peggy Panda, Michael Kelsay—Enterprise Data Architecture Team
- Linda Kosmin—Library and Research Services
- Ed Sewall—Editor
- Wendy Masri—Editor
- Mike Shao—Center of Excellence, Project Scientist
- Sugi Sorensen—Web and Project Information System development
- Bill Spuck—Technology Programs
- Andrea Stein—Technical Information
- Randy Taylor—NASA 7120.5A Requirements
- Dave Werntz—IBS, NBS
- Rebecca Wheeler—Proposals
- Omar El-Sawy, Ann Majchrzak, Alexander Hars—Marshall School of Business, USC
- Paul Gray—Drucker School of Management, Claremont Graduate University

In addition, we received support from the following program, project, and line organizations:

- TMOD
- Space Science Flight Experiments (79X)
- Space Science Flight Projects (74X)
- Earth Science Data Systems (738)
- ESD Leadership Council

TABLE OF CONTENTS

EXECUTIVE SUMMARY

What "Knowledge Management" Means

Definitions of knowledge management (KM) abound in the management literature and consulting communities. Most of these definitions focus on the near-term benefits of improved management of collective business expertise in an enterprise. In this context, knowledge management is important because it increases the ability of an enterprise to find and act on the information that its employees already know. Staying good at this type of knowledge management is said to provide one of the only sustainable competitive advantages in the modern global economy.

There is another, less common definition of knowledge management that adds to the above the care and feeding of an organization's abilities to produce and use new knowledge. Under this definition, the scope of knowledge management also includes efforts to nurture the people, processes, and tools that enable an enterprise to invent new business expertise. Examples of "new business expertise" at JPL include aerobraking to adjust spacecraft orbits and precise determination of Earth satellite orbits via the GPS constellation. If this dimension of knowledge management can be accomplished, an enterprise has even more of a sustainable advantage.

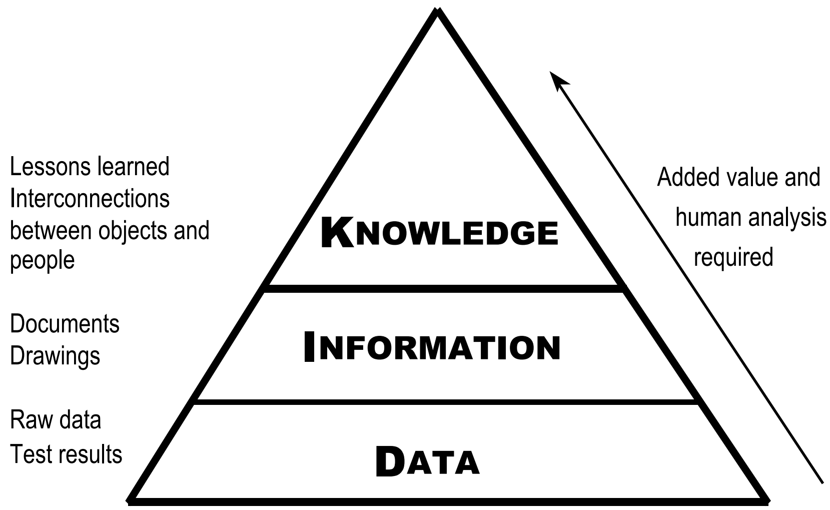
For this study, the Knowledge Management Study Team started with the useful (and reusable) definition produced by the JPL Knowledge Management Workshop in August 1998:

> *Knowledge management is the process of making relevant information available quickly and easily for people to use productively. Information management is the process that focuses on the acquisition, arrangement, storage, retrieval, and use of information to produce knowledge. Any successful knowledge management program will address, as a minimum, the following concepts and issues:*
>
> - *Reuse*
> - *Sharing*
> - *Relevance as determined by the customer*
> - *Training*
> - *Awareness*
> - *Customer identification and focus*
> - *Funding*

To this, the KM Study Team added another item reflecting the expanded definition of knowledge management:

- *Facilitating the creation of new knowledge*

The JPL KM Workshop made a distinction between knowledge management and information management. The pyramid on the next page helps to illuminate that distinction.



JPL's Aspirations in Knowledge Management

The 1998 JPL Knowledge Management Workshop also produced a vision of the desired future state of knowledge management at JPL. The vision was

- Knowledge management is one of JPL's core competencies
- The knowledge a JPL employee needs to do their job is available
- JPL employees have access to internal and external knowledge sources quickly and from wherever knowledge is needed
- The existence and availability of knowledge, as well as its validity and integrity, is assured
- Mechanisms exist to ensure adherence to legal and regulatory restrictions, and to control access where appropriate
- A knowledge management service base exists
- Knowledge management is economical, efficient, convenient, and easy to do

For the same reasons that we expanded the definition of knowledge management, the KM Study Team added another item to this list

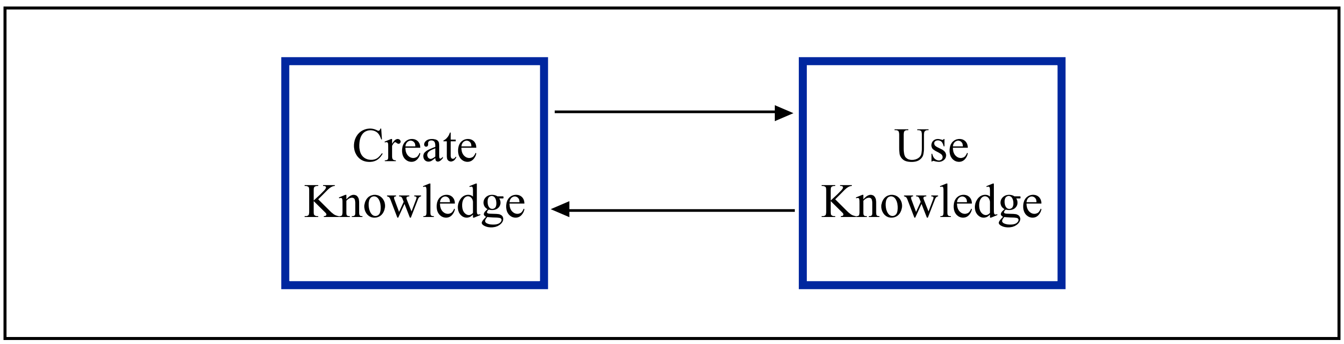
- Knowledge management facilitates the reuse of existing knowledge and the creation of new knowledge

What This Report is About

This report describes some important opportunities to use knowledge management to improve JPL's ability to perform its basic mission of exploring space, and makes some general recommendations on what JPL should do to capture those opportunities. More detailed recommendations will be published in *A Knowledge Management Implementation Plan for JPL*, to be published in February 1999.

The opportunities described here are significant because of their potential to increase the benefits JPL derives from a vital synergy

- JPL and our partners are able to accomplish challenging missions because of what we collectively know
- Accomplishing these missions enriches what we collectively know



The synergy shown in the diagram above is more than just fortuitous—its continuation is essential to the sustained prosperity of a Laboratory whose mission is to do what no one else can do. Because JPL's mission is to continuously expand the frontiers of space exploration and understanding, we have to excel at using and growing our knowledge, and in disseminating that knowledge to our customers.

What This Report is For

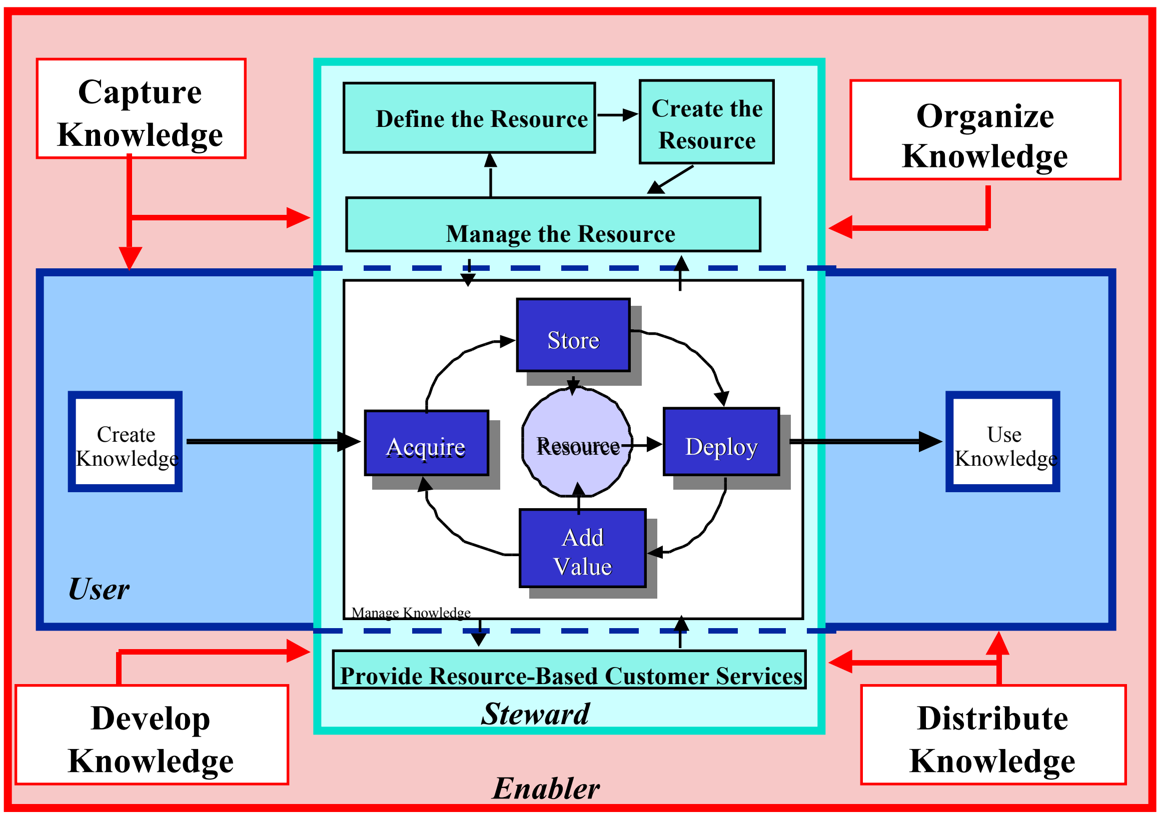
The purpose of this document is to recommend enterprise-scale architectures for processes and services for knowledge management at JPL, and to bring relevance and specificity to those architectures by describing what they are supposed to accomplish or enable. The architectures consist, respectively, of processes and services designed to enable JPL to realize the knowledge management vision stated above. Between the goals and actual results lies a lot of real, focused work. This report supplies both logical continuity and a recommended approach to focus and accomplish that work. The KM Study Team believes that, by following this approach, JPL can transform the abstract goals of knowledge management into real services, provided by real people, to deliver real value to real customers.

Summary of Recommendations

This report makes recommendations in three areas: knowledge management processes, knowledge management services, and near-term knowledge management initiatives.

Knowledge Management Processes

The recommended process architecture for knowledge management at JPL is designed to facilitate the positive interactions between creating and using knowledge. Around that foundation, the KM Study Team designed a process architecture, defining functions and roles to provide effective, sustained services to JPL's knowledge workers, who remain the value-adding mainspring for knowledge management at JPL. The process architecture is displayed in the following chart.



The space limitations of an executive summary preclude explaining this diagram here. Section 3 builds up the architecture by parts, and provides explanations for each part.

Knowledge Management Services

As discussed above, it is through services that the value of improved knowledge management will be delivered to users. Because JPL has been in the knowledge management business for a long time, many of these services presently exist in some form. The intent of these KM Study recommendations is not to displace services that are working well for their users, but to make such services available, at high levels of effectiveness and value, wherever they are needed at JPL. These services are described in the following table and described in more detail in Section 4.

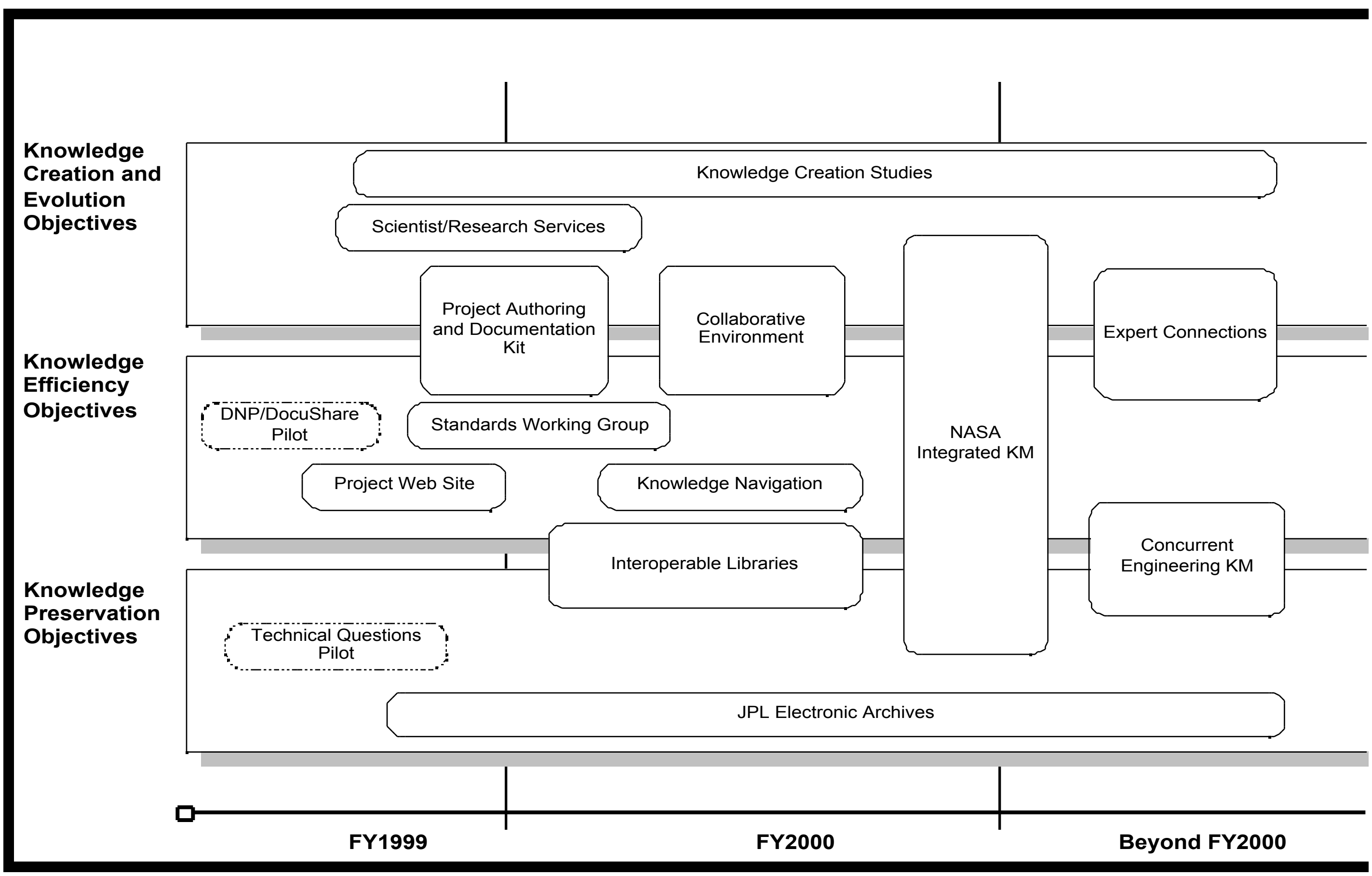
Process	Service	Description
Capture	Resource Development	Provides services for designing and creating knowledge resources including scribing, metric development, and resource assessment
Develop	Authoring	Provides document, engineering design, and software development templates, services, procedures, and standards for authoring new products
	Collaboration	Provides support environments and tools for enabling face-to-face and virtual work teams
	Connection	Provides mechanisms for connecting people, including subject matter experts' directories, interest groups and forums, and electronic discussion groups
Organize	Document and Data Management	Provides services for document and data publishing and retrieval, including metadata standards, document control/versioning, and access control
	Web Site Management	Provides services for Web publishing including metatag standards, design templates and procedures
	Interchange and Conversion	Provides standard data conversion, exchange tools, and processes for core desktop and popular design and authoring applications
	Data Archive	Provides electronic and/or hardcopy repositories and archives of inactive and archival documents, data, and records
	Catalog	Develops core metadata, JPL data categorization standards and taxonomies, and data dictionary
Distribute	Identification	Provides a single user authentication process
	Search, Browse, and Index	Provides functions to browse, search, and index JPL knowledge resources and search across repositories
	Research	Provides expert research services for information retrieval and library subscriptions
	Information Analysis and Mining	Provides tools and techniques for processing and interpreting data
	Workflow	Provides general workflow and electronic forms for automating routine processes
KM		Establishes incentives for contributions to and reuse of knowledge and clarifies rules for knowledge access
	Training	Provides training and communications for using and contributing to the JPL Knowledge Base
	Operations and Maintenance	Provides service and support staff for computer operations, security, metrics collection, and help desk

Near-Term Knowledge Management Initiatives

The knowledge management initiatives recommended here are intended to focus on support to researchers, as well as to projects in various stages of their lifecycles—from proposal through final archive of project data.

These activities were chosen for their importance to the Lab's central work, and for the expected ability of JPL or vendor service providers to deliver the functionality needed. In addition, they provide opportunities for many cross-organizational partnerships at JPL—this is considered a key to the overall success of knowledge management at JPL.

The following chart shows a suggested timing for emphases on each of the 12 early knowledge management tasks identified. The time phasing is approximate, and is meant to show priority and precedence, not accurate beginning dates or duration. The specific initiatives pursued, together with their deliverables, participants, schedules, and budgets will be reported in *A Knowledge Management Implementation Plan for JPL.*



1 Introduction

1.1 Overview

This report discusses some important opportunities to improve JPL's ability to perform its basic mission of exploring space, and what JPL should do to capture those opportunities. The particular opportunities described here have in common their potential to increase the benefits JPL derives from a vital synergy

- JPL and our partners are able to accomplish challenging missions because of what we collectively know
- Accomplishing these missions enriches what we collectively know

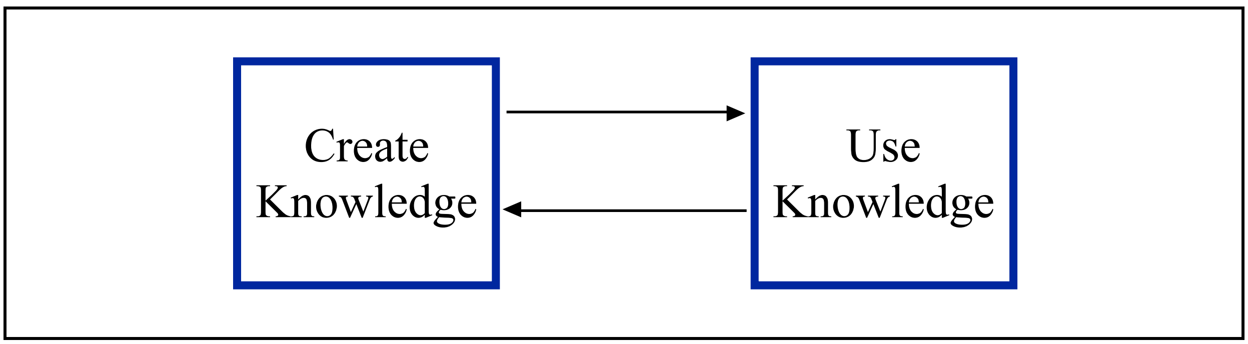


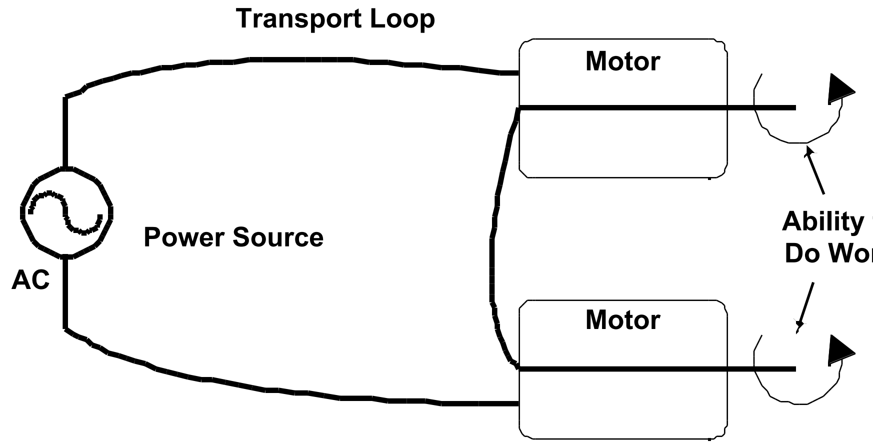
Figure 1-1. The Key Synergy of Knowledge Management

This key synergy, illustrated in Figure 1-1, is more than just fortuitous—its continuation is essential to the sustained prosperity of a Laboratory whose mission is to do what no one else can do. Because JPL's mission is to continuously expand the frontiers of space, we *have to* excel at using and growing our knowledge.

The opportunities in this report also exhibit many differences. A key difference is that between improvements directed at increasing the efficiency of knowledge "traffic" around this key synergy, and those directed at the quantity or quality of the knowledge itself. If we compare the key synergy to a highway system, actions like creating good maps, putting up legible street signs, and documenting and enforcing rules of the road are directed at increasing efficiency. Signs and maps make it easier to determine where you are, where you want to go, and how to get there. Rules of the road prescribe particular individual behaviors in a shared system, such as which side of the road you must drive on, how you must behave at intersections, where you may and may not park. While these rules constrain individual creativity, predictability about what other drivers will do makes it possible for the system to carry more traffic at less cost (both to individual motorists and to the highway authority).

This predictability also decreases the time, risk, and frustration required to get from one place to another using the highway system.

Facilitating the routing, movement, and storage of knowledge decreases the "cost" of delivering it. Another analogy helps to understand what this has to do with the key synergy. Compare the key synergy to a simple electrical circuit, consisting of a power source (voltage), connected to a couple of motors through a transport loop (Figure 1-2). The voltage represents the incentives for knowledge creators to have their knowledge used. (The existence and



strength of these incentives is an important topic that will be addressed later. For now, just assume they exist.) The transport loop represents the knowledge delivery system. The motors represent workers able to convert delivered knowledge into useful work.

Figure 1-2. Electrical Circuit

Actions that decrease resistance in the transport loop lead to increases in delivered current, and in the ability of the motors to do work. In knowledge terms, this corresponds to improving the efficiency of the knowledge delivery system. The motivation for investing in improved efficiency of the knowledge delivery system arises from the positive effects such improvements have on the ability of JPL to do work.

Continuing with the circuit analogy, another effective way to increase the work output of the motors is to increase the applied voltage. In knowledge terms, this means increasing the incentives that knowledge creators actually feel to output their knowledge in a form that others can use.

Still another strategy to increase motor output is to improve the efficiency of the motors, so that more of the energy they receive is converted to work. In knowledge terms, this means reducing the energy required to make use of knowledge once it is available. Providing easy conversion of electronic information from native format into the format in which the user will employ it is an example of this strategy. Another example is increasing the receptivity of work producers to knowledge created by someone else. Examples of each of these work-increasing strategies are represented among the opportunities described in this report.

Both of the above analogies make use of physical phenomena to explain some important aspects of the knowledge opportunity. Actually, knowledge is *not* physical, a property that has some very important implications. Unlike most productive resources (e.g. petroleum) knowledge is not consumed by being used. That knowledge can actually be increased through use is the major driver of the key synergy. This unconsumability also means that my use of a piece of knowledge on Tuesday morning has no impact on your ability to use the same knowledge at the same time. Because I have not consumed the knowledge (and perhaps even enhanced it), you and I can use it again on Wednesday morning.

Among the efficiency opportunities in this report, you will see many that are familiar from traditional information management. Examples are making information easier to find, reducing multiple entries of information, and eliminating the potential for error that arises from storing the same information in multiple places. Much of the overuse of the term "knowledge management" consists of re-labeling such well-known opportunities as if they (and the products and consultants who specialize in them) were something new and different. It is important to recognize that while the absence of "newness" in these opportunities may inspire skepticism, it does not make them any less important. *X* amount of synergy is equally valuable to JPL, whether it arises from new knowledge or better use of existing knowledge.

The purpose of the recommendations in this document is to make use of opportunities to improve how knowledge is produced and used at JPL, whether or not those opportunities are "new." The KM Study Team considers it an advantage that the enterprise knowledge management services recommended make use of many capabilities that either already exist or are already in development.

1.2 Definition of Knowledge Management

Definitions abound in the literature for knowledge management. Most of the leading definitions focus on the practice and process of managing collective business expertise in an enterprise. It is the ability of an organization to find and act on the information that its employees already know. For the purposes of this study, the team used the definition set forth during the Knowledge Management Workshop:

> *Knowledge management is the process of making relevant information available quickly and easily for people to use productively. Information management is the process that focuses on the acquisition, arrangement, storage, retrieval, and use of information to produce knowledge. Any successful knowledge management program will address, as a minimum, the following concepts and issues:*
>
> - *Reuse*
> - *Sharing*
> - *Relevance as determined by the customer*

- *Training*
- *Awareness*
- *Customer identification and focus*
- *Funding*
- *Facilitating the creation of new knowledge*

1.3 Document Purpose

The purpose of this document is to recommend an enterprise Knowledge Management System and Process Architecture for JPL. The knowledge management system and process architecture is defined as a set of services to more effectively manage JPL's most important resource—it's corporate knowledge.

JPL has always been in the "knowledge management" business. As discussed above, our success as an institution is largely the result of our talented and creative workforce and their ability to do what "nobody has done before" in our prime mission to explore deep space. As we adapt to the faster, better, cheaper implementation model, sharing knowledge has become more of a challenge. Some of today's realities that impact our effective use, reuse, and sharing of knowledge include

- Completing just the contracted deliverables is challenging under faster, better, cheaper resources and schedules
- Project development teams are often virtual, and include industry, academic, and/or foreign partners
- Employees support multiple projects at the same time
- Employee turnover is higher than before
- New information and technologies are coming to fast to absorb into work processes

These knowledge management services are specifically designed to complement and integrate with existing infrastructure capabilities like the Enterprise Information System (EIS), the Information Business Solutions (IBS), and Technical Information and Library Services, so that investments the institution has made over the past decade continue on a natural evolutionary path to improve the way JPL manages it's core business.

All of the proposed services contribute in some way to more effective knowledge management at JPL. In several cases, the proposed services are similar to those suggested as potential "value-added" capabilities to be added at a later date on top of the EIS service base. Others may be considered extensions or changes to existing processes or services provided by JPL organizations, vendors, or outsourced operations and should be considered for implementation by the appropriate organizations.

This document is intended to provide the basis for developing a phased implementation plan to develop a knowledge management system and a set of initial knowledge bases for JPL. The document is organized to follow the logical progression the KM Study Team used to define the overall knowledge management architecture. Section 1 provides a general introduction and

JPL's vision and strategic drivers for knowledge management. This is followed by an initial set of high level requirements in Section 2. Sections 3 through 5 provide the core of the study results with a knowledge management process architecture, a set of knowledge management services, and a knowledge management information systems architecture. Priorities and recommendations for implementation are provided in Section 6, followed by examples of knowledge resources from primary areas and process domains at JPL in Section 7. Several appendices are provided that include detailed work materials used to establish the recommended architecture and implementation priorities.

1.4 Scope

The scope of this document is the JPL knowledge management system, an enterprise level system designed to provide the foundation for the creation, storage, dissemination, and reuse of JPL's knowledge. It is an "enterprise" system, since it is content independent, and therefore provides the building blocks necessary to support the diversity of data, information, and knowledge types that support our work at JPL. Components of the knowledge management system will be utilized by every JPL employee, partner, and customer.

The knowledge management architecture places equal emphasis on knowledge management processes and knowledge management services. As JPL continues its migration to a process-based organization, the redesign and tuning of its business and engineering processes must be dealt with during the architecture and design phases of system development. Without this connection, the institution runs the risk of creating robust and technically sound infrastructure services, but services that fail due to broken processes and interfaces. The knowledge management services provide people to manage and assist with the processes required for effective use of the Knowledge Management services.

This document provides a high level architectural framework for implementation by service providers. It also provides the end user with knowledge management service descriptions and example "use cases" that illustrate the kinds of knowledge-sharing activities these services will enable.

1.5 Study Approach

The knowledge management architecture study consisted of four basic activities: research, organization, architecture recommendations, and implementation recommendations. This document reports the results of the first three. Implementation recommendations are described in a companion report, *A Knowledge Management Implementation Plan for JPL*, published in August 1999.

1.5.1 Research

The research phase addressed both the theory and practice of knowledge management. Managing and growing knowledge assets is receiving significant attention at other high-technology organizations, and from the academic, consulting, and vendor communities. Given the knowledge reuse potential of this outside work, the KM Study Team used consultants, case studies, informal benchmarking, vendor product reviews, and the open literature to determine how other organizations defined the knowledge management problem and how they addressed it. Consultants and the literature helped in understanding what has and has not proven effective. Work was also reviewed for findings or recommendations on what was needed most and what was needed first.

For JPL-specific needs, people from a wide variety of JPL organizations were interviewed to understand their requirements and priorities in finding, creating, and reusing knowledge. The results of the JPL Knowledge Management Workshop held in the spring and summer of 1998 (in which four of the six members of the Knowledge Management Study Team participated), provided a foundation for this report.

1.5.2 Organization

To succeed in designing institutional knowledge management solutions that would be efficient and effective across a wide range of locally specific needs, it was essential to find common elements within that variety of needs. Thus the Organization phase consisted of sorting through the large quantity of material obtained in the research phase, and reducing that information to common themes and logical groupings of needs.

1.5.3 Architectural Recommendations

To meet the needs identified and grouped in the organization phase, a set of knowledge management services was derived. Casting the proposed institutional solutions as "services" emphasized some key issues

- Definition and scope of the service, defining what is done by the service, and what remains the responsibility of the local user
- Service ownership, which includes both
  - Provision responsibility for maintaining and evolving the talent and tools to support continued service provision (the "service base")
  - Process responsibility for ensuring that the service continues to meet both customer needs and external requirements and constraints
- Integrating the various elements of a service into an accessible and accountable point of contact, so that users can get expert assistance

After defining knowledge management services corresponding to the grouped needs, the remaining work of this phase was determining priorities, given the scarcity of implementation resources, for implementing the services. Priorities were determined by striking a balance among multiple, and sometimes competing, objectives. The knowledge management services selected for early implementation needed to be

- **Relevant**: Meeting an identified need should provide, within some reasonable time frame, a clear and important contribution to JPL's ability to pursue its core mission of exploring space
- **Sustainable**: The processes, systems, and technologies needed to manage knowledge should not just be "dropped on" JPL. Instead, the knowledge management architecture should provide for a sustained capability to deliver the benefits of these processes to customers focused on their own jobs.
- **People-Centric**: While technology is an important ingredient to effective knowledge management, humans are the essential ingredient, both for knowledge production and for knowledge use. Any effective knowledge management process or system must add value that, in the minds of its users, exceeds the cost of its use.
- **Feasible**: Implementing the knowledge management architecture requires funds and specific skills, both of which are scarce. The recommended knowledge management priorities of this report were made after considering the above objectives, as well as considerations of technical prerequisites (capabilities needed early to support subsequent initiatives).

1.5.4 Guiding Principles

The Study Team followed a set of guiding principles that forged the Team's commitment and helped evolve the architecture. These principles provide the foundation for the architecture. Caution should be practiced in making changes to the architecture that violate these principles.

- Emphasize that knowledge is a Lab resource
  - Ensure knowledge is easily accessible and shared by all
  - Maintain essential knowledge required to meet the Lab's needs
  - Ensure that knowledge creation and capture provide for secondary use
- Maintain and improve the quality, integrity, timeliness, and access to knowledge
  - Allow for data to be maintained separate from the supporting application
  - Define data consistently across the Lab
  - Provide enterprise-strength management tools to information owners
  - Minimize the duplication of information
- Minimize Lab-wide costs
  - Provide a flexible, expandable architecture
  - Build only what is necessary to complete a capability
    - Combine existing capabilities
    - Upgrade existing services where necessary to meet knowledge management requirements
  - Provide enterprise services when appropriate
  - Provide enterprise support for enterprise services
    - Interoperability solutions
    - Professional service base
    - Migration tools
    - Application support and training

- Application refreshment
- Standardize only what is necessary to hold things together
  - Use of services is optional
- Meet JPL's legal, policy, and security needs
- Measure the quality of the knowledge, system, and supporting services

1.6 Knowledge Architecture

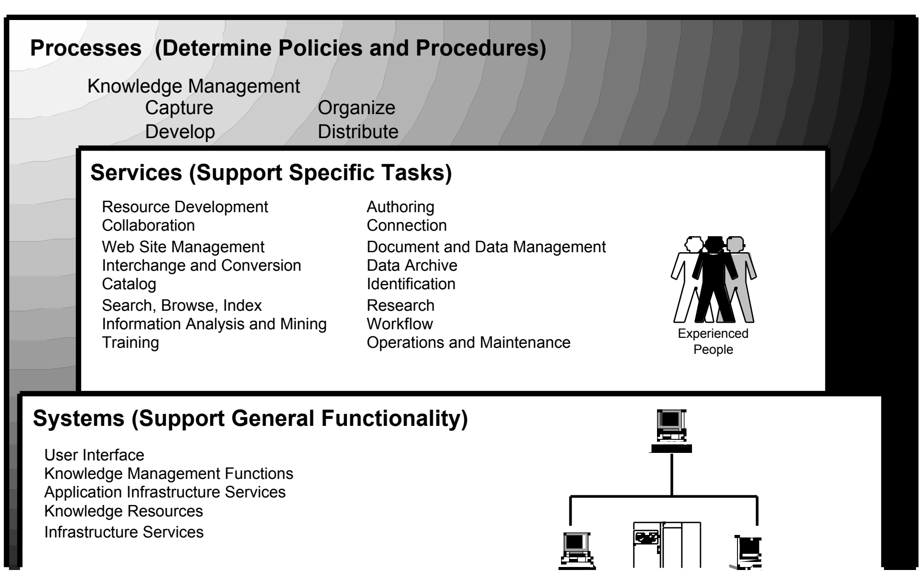
In looking at how other companies are successfully managing knowledge, it is clearly necessary to have a unified approach within an organization on how knowledge is captured, developed, organized, and distributed. As a result, we recommend an architectural approach to JPL's knowledge management that incorporates processes, services, and information technology components.

The central tenet of this knowledge architecture focuses on helping people do their work more effectively. The architecture is centered on services that offer people expertise, one-on-one support, templates, and tools. The three components of the architecture are

- **Process**: Centered around JPL's core business processes, knowledge management processes guide and govern how people share knowledge through the use of policies and procedures
- **Services**: Services are staffed by people that can assist projects and individuals in developing, capturing, organizing, and distributing their knowledge
- **System**: The knowledge management system is a federated arrangement of both centralized and locally controlled systems, tools, and technologies that follow similar procedures and standards for interoperability and information exchange

For example, if a JPL employee were trying to convey knowledge to someone, they would follow a process, possibly getting help from a service provider, and probably using a computer system to send the information. In this case, the employee would interact with all components of the architecture.

Figure 1-3 illustrates the integrated nature of these components, where services are how



people get support in their day-to-day work, systems provide an infrastructure for communicating and capturing information, and processes provide a background of governing policies and procedures. The sections that follow detail the specific processes, services, and system components that are related directly to knowledge management.

Figure 1-3. An Integrated View of the Knowledge Architecture

2 General Requirements

The purpose of requirements gathering during the architecture phase was to provide a mechanism to help identify themes that appear frequently. This then indicates a need for knowledge management services or processes to satisfy the requests for specific support. Requirements presented here are therefore at a high level and are not intended for detailed traceablility. They can however be allocated to one or more knowledge management services or processes where they can be expanded and extended for detailed design, implementation, and testing.

The following are some fundamental knowledge management requirements grouped into three categories of primary business priorities:

Knowledge Efficiency—Productivity Improvements

- Reduce the time to find the right knowledge for the job
- Reduce complexity of work processes (remove irritants)
- Provide intuitive and Web-based tools that require minimal training for effective use
- Provide an integrated view of JPL knowledge

Knowledge Creation and Evolution

- Provide easy mechanisms for knowledge capture and access
- Enable effective team collaboration
- Enable knowledge reuse
- Provide targeted (filtered) knowledge relevant to individual needs

Knowledge Preservation

- Automate compliance with rules, regulations, and policies
- Provide lessons-learned and best practices knowledge bases
- Verify content and provide context to knowledge bases

2.1 Methodology

A standard waterfall approach to requirements gathering was not used for this study. As Figure 2-1 illustrates, requirements were obtained from a variety of sources.

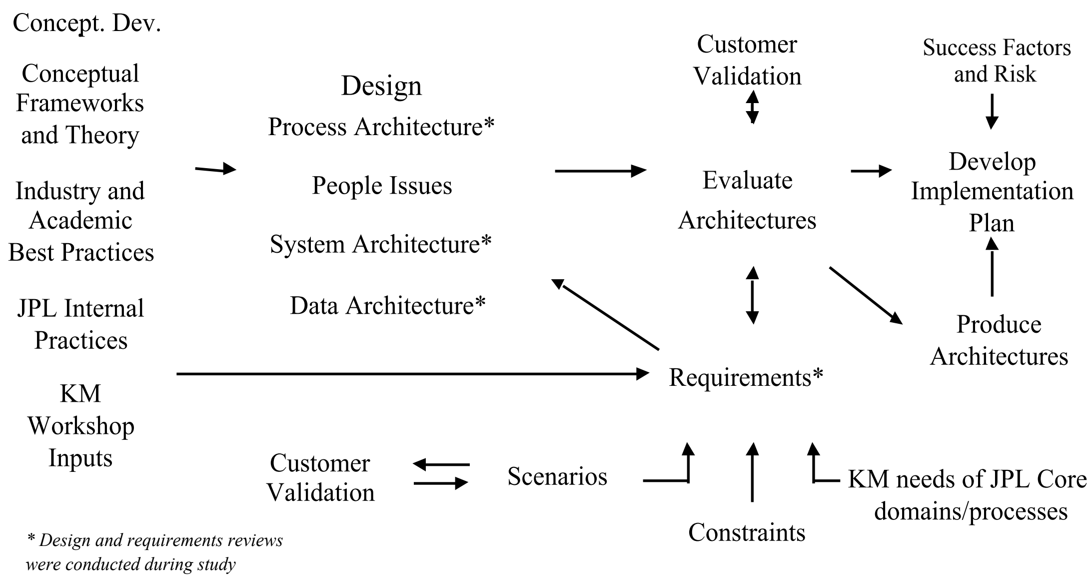


Figure 2-1. Knowledge Management Study Approach

2.1.1 *Conceptual Development*

A large number of ideas or possibilities were generated during the Knowledge Management Workshop. Many of these could be considered potential requirements on a knowledge management system. Those items given higher ranks by the collective group were reworded and cast as potential requirements. In addition, research on knowledge management conceptual theory and frameworks, industry and academic best practices, benchmarks, and JPL internal practices all yielded additional insight into requirements typically addressed by a knowledge management system.

2.1.2 Scenarios, Interviews, and Reviews

Interviews and reviews were conducted with potential knowledge management users and providers during the study period. These discussions provided additional requirements, and typically focused on the need for increased productivity. A primary concern was the lack of time to devote to *any* training to use new technologies. Suggestions were received by the group in a variety of forms and consolidated into comments

2.1.3 Constraints

A number of policies and constraints are at odds with the desired state of creating a knowledge-sharing culture. The need to comply with ISO 9001, NASA’s new guidelines for Program and Projects (NPG 7120.5A), and the tightening of security for document and data release to non-JPL partners provides additional requirements.

2.1.4 Knowledge Management Needs of Core Domains and Processes

Representatives from each of the top seven process domains at JPL were consulted during the study for identifying knowledge management needs and priorities within each of their processes. While there may be different priorities placed on various knowledge management services in each of these areas, there was consistent support for the primary knowledge management requirements. Emphasis was given to requirements to enable knowledge creation and dissemination within and between processes.

2.1.5 Architecture Evaluations

In evaluating system and data architecture components for knowledge management, a variety of interface requirements to existing systems and services were identified. Appendix G provides a detailed list of requirements identified and their allocation to proposed knowledge management services.

2.2 Assumptions

The knowledge management architecture makes the following key assumptions about the user community at JPL. These are technically not requirements on the knowledge management system or processes, but need to be explicitly stated. Without agreement on these underlying principles, the overall knowledge management architecture will not work.

- **Web Literate**: JPL users of the knowledge management system will be trained and comfortable with using a Web browser in their daily work processes. This includes the ability, at a minimum, to update browser configurations with the aid of a human and/or help file to add or change MIME-type settings, bookmark pages, use file upload features of the browser software, and utilize Web-based search engines.

- **Web Enabled**: JPL users of the knowledge management system will have desktop and/or mobile computer systems available for their use. These computers will provide standard browser software installed and configured for access to JPL standard file types as specified by the knowledge management system. Browsers will be configured with core product applications installed and configured for launching native file formats when required.

- **Security Aware**: JPL users of the knowledge management system will be trained on security requirements for different types of JPL data, designs, software, and documentation. Users will be trained on the proper use of security features of knowledge management systems and the processes and procedures for obtaining clearances for publishing knowledge to non-JPL users. Users will be held responsible for violations to basic security principles and improper release of protected information.

3 Process Architecture

The first component of the recommended knowledge architecture involves processes. As a process-based organization, JPL's policies and procedures governing knowledge management will be derived from processes.

Knowledge management should exist as a single process within the Provide Enabling Services (PES) subdomain and have the following subprocesses:

- **Knowledge Management (process)**
  Ensuring that the subprocesses and supporting services and systems work together to achieve the knowledge management objectives within JPL's constraints
  - **Capture (subprocess)**
    Helping people to articulate their knowledge into a form that is sharable and useful to others at the Laboratory
  - **Develop (subprocess)**
    Selecting and refining material to increase its value for users
  - **Organize (subprocess)**
    Providing "building blocks" for storing knowledge, defining standards, and standardizing the way knowledge is described to make knowledge easier to find and retrieve
  - **Distribute (subprocess)**
    Facilitating how people gain access to material

The following sections describe the evolution of these processes and the detailed descriptions of each. At its core, knowledge management processes exist from three perspectives:

- **Users**: people who create and use knowledge as part of doing work
- **Stewards**: people who create knowledge resources to help themselves and others do work
- **Enablers**: people who create re-usable structures, processes, and services that help stewards to develop their resources and users to use them.

The process architecture presented in this section addresses each of these perspectives and how they relate to one another.

3.1 Perspectives

3.1.1 User Perspective

A *user* is someone who, in the process of doing his or her job, creates new knowledge and makes use of existing knowledge. The key synergy (Figure 1-1) shows this continuous cycle of creating and using knowledge. This cycle is supported by processes that enable the exchange of information. Knowledge management at the user level consists of acquiring, storing, deploying, and adding value to knowledge. Acquiring is the means of codifying the information or inputting it into a shareable form. Storing maintains the information in usable formats. Deploying provides the output mechanism and ways in which the information can both proactively and reactively be provided to users. Adding value provides the means to improve the information, making it more valuable to the users. *The user has the knowledge; the knowledge management process allows the user to structure and share it*. In Figure 3-1[1], the connections between creating and using are expanded to explicitly show the knowledge management process.

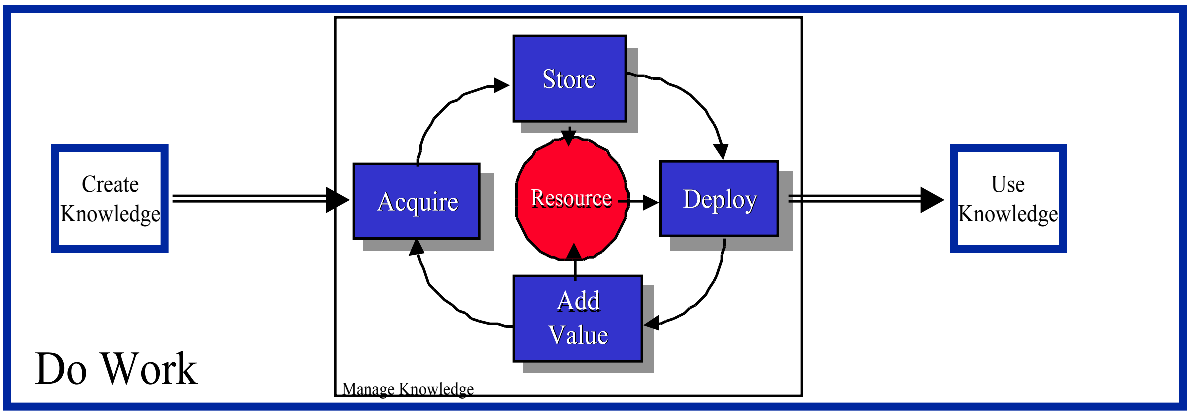


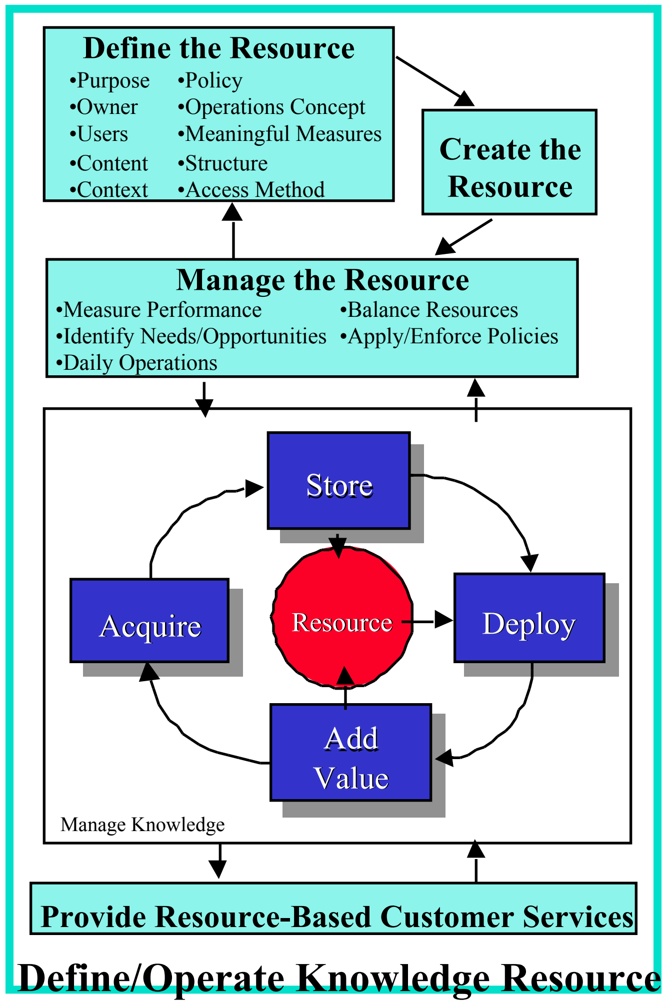
Figure 3-1. Expanded User Perspective of Knowledge Management

3.1.2 Steward Perspective

Stewards are people who create knowledge resources to help themselves or others do work. There are at least two forms of stewardship. First, individuals on Lab, in the course of doing their jobs, build up personal knowledge and expertise that can evolve into a shareable resource. This knowledge begins as a set of personal information or experience that the individual calls on to do his job. As this base of knowledge grows, the person develops ways of organizing, storing, and enhancing the information locally to make it easier for him or her to use it in future efforts. Eventually, this knowledge may become useful to others in performing their jobs, if it is available to them. Stewardship of knowledge developed through personal expertise is an individual responsibility and an organizational necessity.

The second form of stewardship is via process ownership. By definition, process owners are responsible for the knowledge needed to use and operate their processes. Every process owner will be defining and operating knowledge resources specific to their process. One way of improving a process is to identify best practices and collected expertise, and embed them into the process. By doing this, the best and most productive ideas can be made available to all

[1] The Acquire-Store-Deploy-Add Value cycle is adapted from Ernst and Young.



who use the process. Stewardship of knowledge developed specifically to support a process is required of all process owners.

Figure 3-2. Stewardship Perspective of Knowledge Management

The knowledge management process from the stewardship perspective is shown in Figure 3-2. It address both people issues and knowledge captured in electronic or other forms. The stewardship processes of knowledge management are resource-centric. They center around the creation, operation, and use of a specific resource, such as a database, collection, file system, group of specialists, knowledge base, or Web site. The processes that the steward performs are listed in the following sections.

3.1.2.1 Define the Resource

The first step is to define the resource from the perspectives of the users and the stewards. There are several dimensions to this definition: purpose, owner, users, content, context, policy, operations concept, meaningful measures, structure, and access methods. By addressing each of these dimensions, the steward clearly indicates why the resource exists, how users will use it, what the content is, when it's appropriate to use it, the level of reliability of the information, how it will change and grow over its lifecycle, and how the steward will know if it is indeed useful.

3.1.2.2 Create the Resource

Creating a resource, in general, will involve the development and integration of information system components to provide the required functionality. The information system provides the structure and means of putting information in and retrieving information from the resource. The second part of creating a resource is populating it with the appropriate data, information, and knowledge. Both of these steps must be performed in order to have an operational system.

3.1.2.3 Manage the Resource

Once the resource has been created, the steward is then responsible for managing the resource. This involves functions such as measuring performance, identifying needs and opportunities for improvement, obtaining and distributing resources, applying and enforcing policies, and conducting daily operations. Part of managing the resource is to address feedback from customers of and contributors to the resource.

3.1.2.4 Provide Resource-Based Customer Services

For some resources, there are additional services that the steward can provide that help people use or contribute to the resource. For example, if the resource requires information to be collected in special formats, a resource-based service might put customer data into that format. Other services might provide assistance with finding and retrieving information from the resource, e.g., a reference librarian assisting in the use of a library. Part of managing the resource is to identify what types of services would improve usefulness to the spectrum of customers using the resource.

3.1.3 Enabler Perspective

Enablers are the people who create re-usable structures, processes, and services that help stewards to develop their resources and users to use them. While stewards (either individuals or process owners) ensure that the knowledge is available to support the *primary use* activities for which it was created in the first place, enablers are responsible for identifying and supporting *secondary use* of the knowledge. Secondary use refers collectively to all the other people who may benefit from having access to knowledge created by someone else for a specific purpose. For example, our flight projects create large numbers of documents and other products that they need in order to build the spacecraft and operate the mission. The primary use of these products is by the flight project that created them. However, many of those products could be useful to other flight projects, either as examples or reusable products. The reuse of one project's documents by another project is one example of secondary use.

The enablers are responsible for identifying effective and efficient means of supporting sharing, reuse, and compliance to constraints across all of the knowledge resources on Lab. Technology only supports so many ways of structuring, organizing, storing, and deploying information. Part of the enabler process is to identify building blocks that could be used to support the development of a variety of resources. Integrated into the building blocks would be implementations of standards, rules, regulations, and other constraints that could be best addressed once for the institution, rather than many times for completely unique solutions.

The purpose of the enabling processes is to represent the Laboratory's interests in interoperability, cost-effectiveness, ease-of-use, and compliance, while making it as easy and effortless as possible to help the users and stewards concentrate on the content of their knowledge resources. While it isn't always possible to meet institutional needs without placing some requirements on the individual providers and users, the goal of these processes is to minimize the impact. The enabling knowledge management processes are discussed in the following sections. Each enabling process will have a set of associated services. The enabler perspective of knowledge management is shown in Figure 3-3.

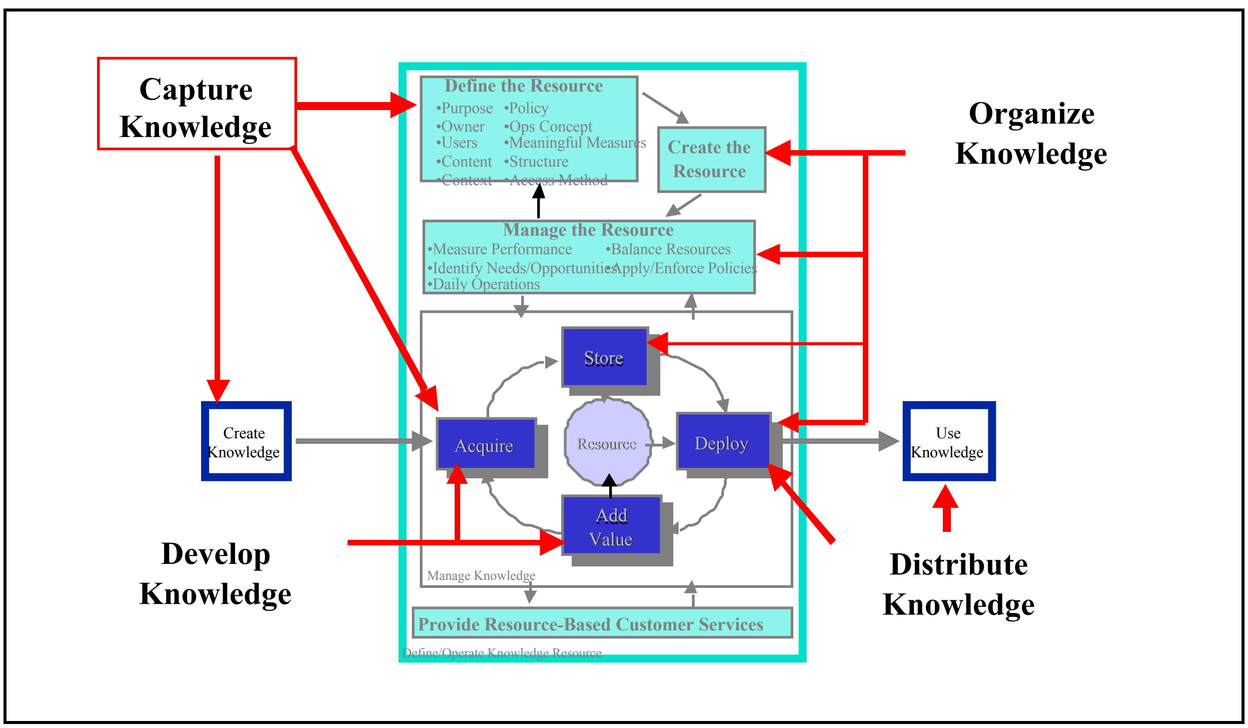


Figure 3-3. Enabler Perspective of Knowledge Management

3.2 Recommended Knowledge Management Processes

3.2.1 Knowledge Management Process

The overall knowledge management process focuses on ensuring the smooth integration of the subprocesses, as well as evaluating the constraints and requirements that need to be met by the entire knowledge management process. The focus of this process is inreach and outreach, such as encouraging use and reuse of knowledge and training people to use the knowledge management system. Services would include training and operations and maintenance.

3.2.2 Capture Knowledge

The capture knowledge process is responsible for helping people to articulate their knowledge into forms that are sharable and useful to others at the Laboratory. This involves identifying potential secondary users, determining the appropriate structure and form in which to capture the knowledge, and defining resources in which to safeguard the knowledge. A significant part of this process is to assist stewards in defining their knowledge resources in a way that meets the needs of both primary and secondary users. It is also responsible for assisting experts in articulating their knowledge, facilitating groups of experts in creating and assessing their combined expertise, and identifying resources that are needed at an institutional level. The Resource Development Service is provided under the capture knowledge process.

3.2.3 Develop Knowledge

The develop knowledge process is responsible for selecting and refining material to increase its value for users. This could include chunking and categorizing knowledge, packaging or re-packaging it to meet specific user needs, and assessing quality and reliability of the knowledge. Services that support this process are authoring, connection, and collaboration.

3.2.4 Organize Knowledge

The key to making knowledge sharable and reusable is to provide structure to it. The organize knowledge process is responsible for providing common, reusable structures in which to store knowledge, defining standards to make the knowledge more broadly available to the JPL community, and providing standard ways of describing the knowledge to make it easier to find, retrieve, and determine applicability. It includes enabling knowledge sharing systems or tools, knowledge bases, navigational devices, user interfaces, taxonomies, filtering knowledge bases and refreshing, deleting, and adding material. There are a number of services associated with the organize knowledge process: document and data management, Web site management, interchange and conversion, archiving, and cataloging.

3.2.5 Distribute Knowledge

The distribute knowledge process is concerned with how people gain access to material. It is responsible for making it easy for people to find information and for providing push and pull technologies. Distributing knowledge goes beyond just making resources available to users. It includes all of the supporting information needed to help users search, find, evaluate, select, and retrieve information. Services included as part of this process are: identification; search, browse, and index; research; information analysis and mining; and workflow.

3.3 Integrated View

The user, steward, and enabler perspectives all need to co-exist and support one another in order for the entire knowledge management process architecture to function effectively. Figure 3-4 shows a detailed, integrated view of the three perspectives. The relationships between the perspectives are simple: users generate and use knowledge; stewards provide needed knowledge to users; and enablers provide systems, tools, and services to help the stewards create resources and users to use them.

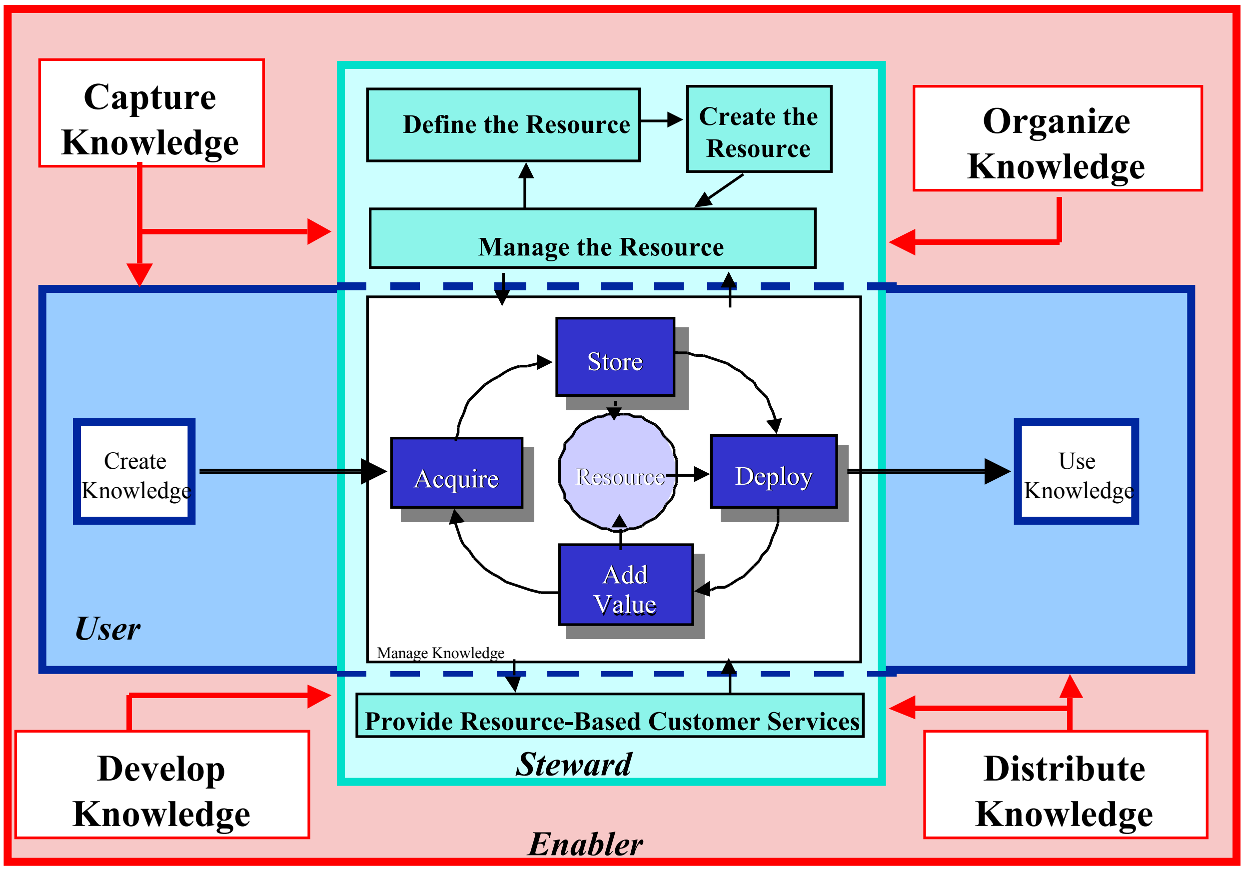


Figure 3-4. An Integrated View of Knowledge Management

3.4 Incentives

A fundamental part of the knowledge management process, common across all levels, is the need to address human motivation issues. The central tenet of knowledge management is the sharing of knowledge. Therefore, it is critical to provide incentives that result in people being willing to share their information—and to make use of information, products, and systems created by others. Inherent in sharing is an understanding of the secondary users of a given set of information. Incentives that reward people for putting in the extra effort required to support secondary user access is critical. Secondary use will require additional effort by all parties—both the creators and the users. Part of the goal of knowledge management is to minimize the additional effort and make it easy, efficient, effective, and valuable to make use of existing knowledge resources when performing a given job and to incorporate the additional context needed to support secondary users. Knowledge management is intended to free up stewards from the routine, system-development aspects of knowledge capture and enable them to concentrate on developing and enhancing the content of knowledge.

The rewards and recognition structure at JPL needs to be changed to acknowledge the contributions of people to JPL's Knowledge Base. This could involve including a knowledge management component to the job descriptions of key personnel such as process owners, group supervisors, and technical principals. It also means providing resources to assist people in converting personal or local knowledge bases into sharable resources.

4 Knowledge Management Services

4.1 Rationale For "Services" Framework

While JPL is continuing to become a "process-based" workplace, many employees still relate their daily work more closely to their roles within a specific project or task they support. Projects and line organizations in turn, look to JPL's institutional services to assist them with infrastructure types of support where it makes sense and when they are available. These services comprise experienced staff in a specialized area, templates developed for JPL projects and programs, and interoperable tools. In addition, the most common model recommended to the KM Study Team was that proposed by the Enterprise Information System (EIS) in 1995 [4]. EIS developed a set of services designed to meet the primary infrastructure requirements at that time. EIS also identified a set of potential "value-added" services that could be added on top of EIS services at some time in the future. Several knowledge management services share some of these earlier visions.

After examining a representative set of potential customer requirements, evaluating constraints, and listening to the most urgent needs of JPL workers, the KM Study Team came up with a list of 16 knowledge management services. The one that most frequently rose to the top of everyone's list for example was an enterprise Document and Data Management Service. Review by our academic consultants, however, and continued review of the knowledge management literature revealed that document and data management was really more appropriately classified as an "infrastructure application" or "information system application" service. These services typically map to primary business objectives for improving efficiency or complying with rules and regulations. While these sorts of services are crucial underpinnings of a knowledge management system or service base, they do not address more visionary knowledge management business objectives for creating and growing true knowledge.

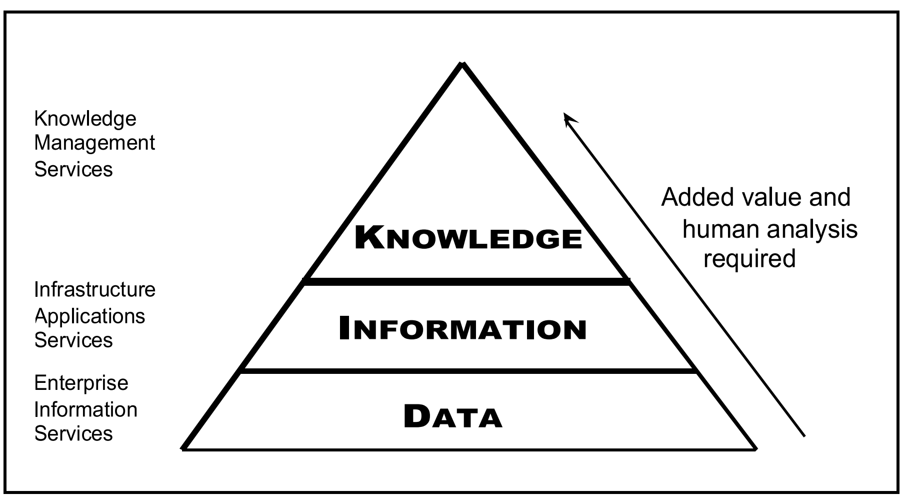


Figure 4-1. Knowledge Management Services and Infrastructure Services

4.2 Services Overview

Table 4-1 provides a summary of recommended knowledge-management related services. Since the knowledge management architecture is designed to be closely coupled with JPL's process-based management structure, each knowledge management service is associated with its primary knowledge management process or subprocess. In addition, one or more of the primary business objectives identified in Section 1 are provided for each service. Finally, each service is identified as primarily a *knowledge management* or a *infrastructure application* service. This service category distinction is designed to aid in implementation planning, and to recognize more general information management services that are needed to support not only knowledge management at JPL, but general best-business practices. In some cases, recommended infrastructure application services exist. They would however benefit from a common direction and set of goals that the knowledge management process and information system architecture can provide.

Table 4-1. Recommended Knowledge Management Services

Knowledge Management		Primary Business Objectives			Service Category	
Process	**Service**	**Efficiency**	**Preservation**	**Growth Creation**	**KM**	**Infrastructure Application**
Capture	Resource Development		X	X	X	
Develop	Authoring	X		X	X	
	Collaboration	X		X	X	
	Connection	X	X	X	X	
Organize	Document and Data Management	X	X			X
	Web Site Management	X	X	X		X
	Interchange and Conversion	X				X
	Data Archive	X	X			X
	Catalog	X	X		X	
Distribute	Identification	X	X			X
	Search, Browse, Index	X				X
	Research	X		X	X	
	Information Analysis and Mining		X	X	X	
	Workflow	X				X
Knowledge Management	Training	X	X	X	X	
	Operations and Maintenance	X	X			X

Each of the services identified in the following sections play an important role in the creation and evolution of an effective knowledge management system and service base at JPL. In some cases, holes in the underlying infrastructure need to be addressed before the knowledge management layer can be added effectively. For each of the services, the following information is provided:

Description	An explanation of the primary service functions.
Category	Identification as primarily a knowledge management service, an infrastructure applications service, or another related service.
Requirements	High-level requirements for iterative review and extension during design and implementation phases.
Specifications and Standards	Potential industry specifications and standards.
Technologies and Products	Some examples of available technologies or products. Not intended to be a comprehensive list.
Recommendation	A target set of interfaces or specifications for the architecture, where applicable (for process-oriented services, a general recommendation). This may also include a discussion.
Rationale	Justifications for the recommendation.

4.3 Knowledge Capture Services

4.3.1 Resource Development

Description—The Resource Development Service has three primary functions

- To identify, prioritize, design, and develop core institutional knowledge resources
- To provide customer support to define and create workgroup- or process-specific knowledge resources
- To support meaningful surveys, metrics, and data collection methods

This service is designed to capture knowledge that is anticipated to be of long-term value and reuse. It is the service most focused on the collection of tacit or soft knowledge, capturing reflective and contextual information learned as a result of the successes and failures of the work done at JPL.

By utilizing the institutional knowledge management Resource Development Service, new resources can be designed to integrate into the larger JPL Knowledge Base. This creates opportunities for the resource to be discovered and accessed beyond the immediate workgroup, and potential reuse opportunities are improved. New resources are typically designed to benefit a workgroup or the entire enterprise and require some additional organization and improved access to be most effective. In order to do this function well, the service provider will need to have knowledge of the resources already available that may relate or apply to the desired knowledge base. This service will provide detailed information of existing knowledge resources available at JPL including attributes, access methods, owner, and custodian of the information.

Metrics provide customer support to define meaningful surveys and data collection tailored to the specific information required.[2] For example, a good survey would include carefully defining the purpose and expected actions from an analysis of customer responses, so the time spent gathering the results is most likely to contribute to meaningful measures. Sample types of metrics include strategic goal metrics such as process performance, efficiency, and cost. Product metrics include delivery specifications and performance criteria. Customer metrics typically reveal satisfaction results, while project metrics may convey status of budget and schedule performance objectives. Metrics are important from a variety of perspectives:

- Customers know that they are being listened to, and that problems they've identified are being addressed
- Process designers and system engineers can help pinpoint parts of a process or a system that needs fixing and understand the root causes of problems
- Technology developers get feedback on newly developed technologies and understand their customers' evolving requirements
- Process owners and managers become more customer focused, and can better develop strategic, tactical and project plans.

Category—Knowledge Management Service

Requirements—The Resource Development Service shall

- Provide institutional knowledge resource prioritization, definition, design, and development
- Provide consulting for customers to design and build new knowledge resources
- Provide training for methods in defining knowledge resources for effective reuse
- Establish a knowledge resource peer review process
- Develop and maintain an institutional lessons-learned knowledge resource
- Provide an interview transcription service
- Provide scribing services
- Publish guidelines on recommended use of the Resource Development Service
- Provide a process for defining metrics
- Provide methods for collecting metric data
- Provide methods for creating actionable custom surveys
- Create a generic metrics knowledge resource for use by projects

[2] A metric is a standard of measurement. A metric becomes useful when the standard of measurement is specifically linked to a JPL or project objective. This enables the collection of measurements that identify where improvements are needed.

Specifications and Standards—None

Technologies and Products—Speech recognition and conversion technologies may be used in the future for interview transcriptions. It is not clear if they are cost effective at this time for knowledge management since the process of culling the information to extract the key lessons or information of importance is difficult to automate.

Recommendation—Start by creating a small number of carefully planned new knowledge resources that are of interest to a wide audience at JPL or to satisfy new requirements in NPG 7120.5A or ISO 9001. Use a low technology approach geared toward more rigorous recording and interviewing of key staff after project successes and failures to record lessons learned. Investigate use of contract court reporters for scribing or interview transcriptions. Build upon metrics expertise already gained by staff in Section 311 as part of efforts to measure Help Desk and DNP process effectiveness.

Rationale—Throughout much of the literature on knowledge management, the primary components of a knowledge management system include collections of well-managed knowledge resources. Creation and continual review of these resources by subject matter experts at an institution is required to ensure they continue to meet their objectives. Continuous addition and refinement of the knowledge resources is required. JPL is required by NPG 7120.5A to supply the NASA Chief Engineer with regular lessons learned. Additionally, one of the most important changes in JPL's move to process-based management is to continuously evaluate and improve on existing processes. In addition, several requirements in NPG 7120.5A require that projects provide careful measurements of milestones throughout their project lifecycle. These metrics must be carefully defined to obtain results that can be used to make continuous improvements.

4.4 Knowledge Development Services

4.4.1 Authoring

Description—This service provides templates, guidelines, and examples for creating documents, engineering, and software products in standard ways, including legal and copyright requirements and security labeling. It develops standards and guidelines designed to facilitate automated content summarization and categorization for later search and retrieval. It also provides validation services for controlled or configuration-managed documents, data, and knowledge resources to improve the reliability and integrity of enterprise knowledge bases. The goal is to automate validation as much as possible, providing or embedding validation tools or macros at the time of knowledge creation. This way responsibility and corrective action can be taken as part of the creation process.

The following kinds of functions would be part of an Authoring Service

- Providing document authoring tools and standards, such as Microsoft templates, XML authoring tools, and DTD development
- Providing document authoring services, such as writing, editing, document design, and validation
- Supporting engineering authoring aids, such as procedures and templates for creating engineering drawings and models, and standard assembly definitions
- Creating software development header templates, naming and design pattern guidelines, make file templates, and build script templates
- Supporting an engineering and software product documentation service with examples of good product design
- Embedding validation tools

Category—Knowledge Management Service

Requirements—The Authoring Service shall

- Provide standards, tools, procedures, examples, and consulting for template development for engineering files, documentation, and software code
- Provide writing, editing, design, validation and documentation services
- Embed standard products and processes into JPL policies and procedures where possible
- Develop content-specific DTDs for high use document types
- Interoperate with the Interchange and Conversion Service
- Provide tools for data quality checks and syntax/spell check tools for enterprise document, data, and knowledge bases
- Provide standards compliance verification and metadata validation of enterprise document, data, and knowledge bases
- Review access controls on document, data, and knowledge bases
- Perform usability checks for public material
- Perform link verification and report broken links

Specifications and Standards—HTML, PDF, RTF, XML, XSL, XLL, DOM, WebDAV, STEP, NASA-STD-2804A

Technologies and Products—Microsoft Office wizards and templates, XML DTDs, SoftQuad Author/Editor, Adobe Framemaker+SGML, Inso DynaText (see http://www.xml.com/xml/pub /authortools for more on XML authoring tools and reviews

Recommendation—Many of these services are already provided by workgroups and existing services. These services should be utilized and extended to cover new document, data, and knowledge bases and geared for supporting the highest use document authoring applications (e.g. Microsoft Office) with migration to XML and related standards for structured documentation. This service should provide instructions and procedures for authoring of engineering models and drawings in standard ways (providing templates wherever possible), pilot template implementations for selected STEP specifications, and automating and embedding validation in knowledge creation processes wherever possible.

Rationale—Projects do not have the time to reinvent structures for standard documentation or products. Additional requirements to satisfy ISO 9001 and NPG 7120.5A have highlighted the importance and need for this service.

4.4.2 Collaboration

Description—This service will provide facilities, tools, and a service base to increase the effectiveness of meetings in creating, capturing, and deploying knowledge. Meetings are an important part of daily work processes at JPL. A great deal of knowledge transfer and idea creation occurs during meetings, where people have the opportunity to work collaboratively. As partnerships expand and interfaces become more distributed and complex, the effectiveness of meetings held “virtually” become a critical factor in achieving success on JPL projects. The sorts of functions a Collaboration Service would provide include

- Integrated voice/video and dataconferencing available to conference rooms and desktops
- Tools for capturing and editing agendas, meeting notes, action items, brainstorming flows, and decision capture with output
- Tools for publishing meeting materials to the Web, e-mail distributions, or related workflow applications (e.g. Action Item Tracker)
- Web-based calendaring and scheduling tools
- Web-accessible knowledge base of conference rooms and collaboration equipment.
- Ability to confidently secure reservations for conference rooms
- Detailed instructions for use of conference room equipment posted at site and available electronically
- Simultaneous computer data conferencing, including shared use of computer applications, to any office or conference room with voice and data connections
- Shared whiteboards for real-time markup during conferencing using computer-driven projectors or desktop tools

Category—Knowledge Management Service.

Requirements—The Collaboration Service shall

- Develop and publish standards for desktop and group conference room equipment for various levels of conferencing
- Provide installation services for standard recommended networked equipment

- Provide instructions for operation of recommended equipment and application software
- Provide integrated voice/video and dataconferencing from desktops and conference rooms
- Provide tools for decision capture, action item assignment, and brainstorming
- Provide shared whiteboard capabilities in conference rooms
- Provide Web-based scheduling tools for collaboration and conference sessions

Specifications and Standards—T.120, H.323

Technologies and Products—DataBeam, DataBeam FarSite, Meeting Place, Latitude Communication server, Microsoft Net Meeting, Intel Proshare, Future Labs TALKShow, PictureTel LiveShare

Recommendation—Early tasks for this service should focus on completing the deployment of and providing a service base for Latitude Communications' integrated package for voice and dataconferencing. Continue prototyping and evaluation of new technologies for expanded collaboration capabilities.

Rationale—Collaboration technologies are considered a key component of effective and relevant knowledge management endeavors. In addition, project engineering and management staff should not have to spend their time planning and arranging for facilities and associated equipment. Emphasis should be given to low-cost voice- and dataconferencing in the majority of medium to large conference rooms. The Collaborative Engineering team is already doing this work at a pilot level with several projects and in conjunction with other NASA Centers. As these pilots prove their usefulness to the general Laboratory, these should be brought into the enterprise service base.

4.4.3 Connection

Description—The primary goal of the Connection Service is to provide a variety of methods for users to share ideas, both physically and virtually. It is primarily a people-centric service, and a key one for facilitating true knowledge transfer rather than just information management functions. A Connection Service could assist in developing subject matter experts' directories for people or organizations, such as Centers of Excellence or Technical Communities. It can provide specific subject interest groups utilizing forums for connecting people with similar technical interests, including electronic threaded discussion groups, newsgroups, chat rooms, and service and training for moderators.

Category—Knowledge Management Service.

Requirements—The Connection Service shall

- Identify and assist the development of key subject matter experts' directories

- Provide alternative ways for interest groups to share their ideas in both virtual and physical settings
- Work with Human Resources to identify policies that can help ensure knowledge transfer from key experts at JPL

Specifications and Standards—NNTP, SMTP, IRC

Technologies and Products—Hypermail, OpenText Livelink, News servers, Chat servers

Recommendation—This service should focus on designing and populating one or more experts' databases. Extend use of traditional mechanisms such as brown bag seminars and noon-time speakers. Work with Human Resources to establish incentives for knowledge sharing. An example of this is for all "principle" level employees to present talks and/or mentor colleagues to ensure their knowledge gets passed along. Another example is that key specialists in engineering or science disciplines be available to support multiple projects a certain percentage of time. This service should partner with EIS Messaging for extending service base beyond newsgroups to threaded discussion groups, and possibly chat rooms for selected uses.

Rationale—Most of the research in the knowledge management area points to the value of creating and maintaining simple directories of people as subject matter experts. Much of the current literature on knowledge management focuses on the need for providing a variety of ways for people to communicate and share their knowledge.

4.5 Knowledge Organization Services

4.5.1 Document and Data Management

Description—The Document and Data Management Service provides publishing, versioning, configuration management, and archiving of documents and data through the life cycle. Documents may include policies, procedures, project documents, program documents, and administrative documents. Data types include a variety of object types such as design models, drawings, photos, videos, and software. This service should support information in hardcopy, electronic, and other media formats. Standard functions provided by this service include

- Electronic repositories for documents and data in various stages of their lifecycle and under various control levels
- Document and data clearance services for release of information to customers, partners, and the public
- Document and data configuration management, including approvals, check in/out, revision control, and validation
- Relationships to failure reports, change requests, change control board activities
- Audit trails

- Document and data associations—providing relationships between hardware and or software assemblies or builds, policies and standards, other cross-references
- Hardware development, test, release, red-line, and bill of material management
- Software configuration and build management
- Configuration management planning and procedure development
- Generation of project Master Controlled Data Lists (MCDLs) required by ISO 9001
- Data management capabilities for Ground Data Systems development

It should be noted that the scope of this service is geared toward the management of documents and data products generated during the proposal, design, and development phases of a project. The management of downlinked telemetry data and derivative science products was not researched, although there is no reason why the service described here could not be used for these data also.

Category—Infrastructure Applications Service
Requirements—The Document and Data Management Service shall

- Provide document and data publishing capabilities for public, partner, enterprise, or workgroup (e.g. Project or task) dissemination
- Provide for interoperability between distributed document and data repositories
- Satisfy ISO 9001 compliance for Element 5, document and data control
- Require standard metadata assignment for each object type
- Embed access rights labeling in the publishing process
- Provide Web-based internet and intranet access to documents and data
- Provide an application programming interface (API) for server-to-server document and data exchange
- Utilize EIS unique user names and principles
- Integrate with the EIS Authentication and Authorization services, where possible
- Interoperate with EIS Messaging Service for notification, where possible
- Utilize the EIS File Service, where possible
- Utilize the knowledge management Identification Service
- Provide data quality checks of enterprise documents and data
- Provide syntax/spell check tools for enterprise documents and data
- Provide metadata validation for enterprise documents and data
- Provide review processes for access controls on document, data, and knowledge bases
- Perform routine virus scans on documents and data

Specifications and Standards—WebDAV, ODMA, DMA (Figure 4-2)

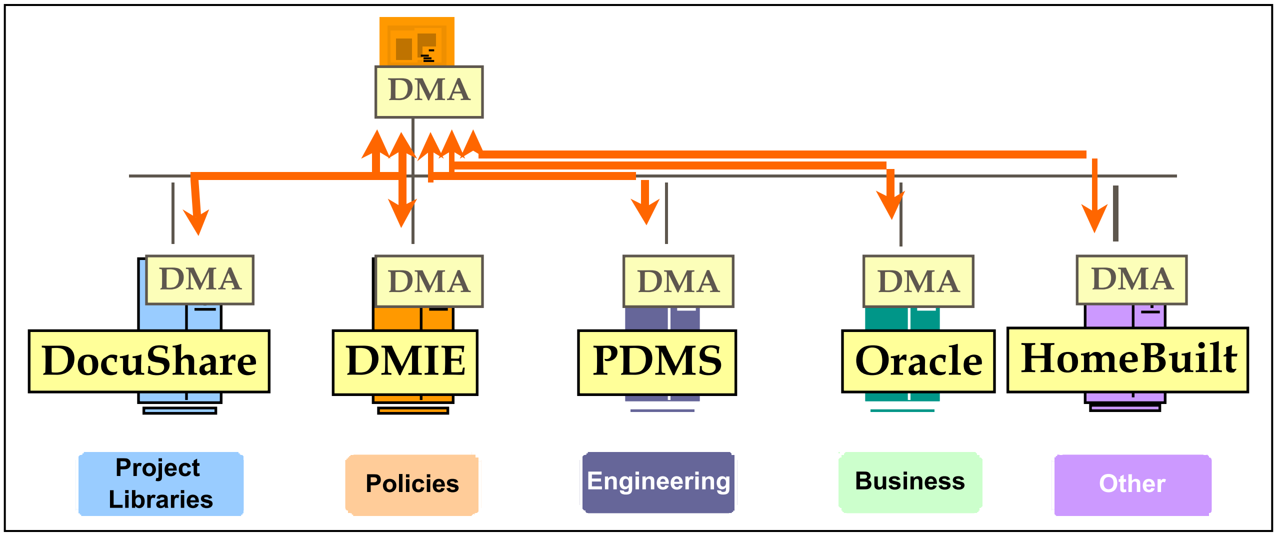


Figure 4-2. Sample Use of Emerging Document Exchange Standards Across Disciplines and Applications

Technologies and Products—JPL utilizes many custom-developed and some COTS products to manage the diverse types of document and data products it produces. There are large numbers of COTS packages available for supporting many of the traditional data types JPL needs to manage. These range from very large integrated systems such as Lotus Notes, to lightweight and flexible document management systems such as Xerox's DocuShare. Some samples of vendors are

- Document Management: Documentum, PCDocs, IBM/Lotus, OpenText Livelink, FileNet, Xerox DocuShare, Astoria
- Product Data Management: Sherpa, WindChill, Oracle (manufacturing module)
- Software Configuration Management: Platinum CCC/Harvest, SCCS, RCS
- Requirements Management: Doors

Recommendation—This service should look to consolidate existing systems from many to a few interoperable systems. Provide continuity by establishing core metadata standards and mechanisms for cross-repository searching. Continue OpenText Livelink pilot and evaluate deployment of single enterprise-wide document and/or product data management system. Include representatives from TMOD, Raytheon (Science Data Systems), and CalTech (IPAC) to determine potential for utilizing this service for their data management needs. This service may be a logical "value-added" application extension to the Enterprise Information System.

Rationale—As employees work on individual projects for shorter periods and often support more than one project at a time, there is no time to build or learn new systems for managing traditional types of information. Project funds should not need to be spent on Document and Data Management Services that the infrastructure could already provide (for at least a limited selection of products). Large and small projects have different needs and partner relationships, so flexibility will continue to be a requirement. A single large-scale deployment of an integrated system has potential advantages, but also has additional risks in higher costs, longer deployment times, heavier training requirements, and large-scale culture changes. Consultant discussions have indicated that incremental changes to infrastructure services tend to be accepted more readily, especially in engineering organizations where diversity and creativity is highly valued.

4.5.2 Web Site Management

Description—The Web Site Management Service provides design, development, publication, configuration management, archive, and maintenance services and tools for internal, project, and public Web sites. This service would provide a one-stop shop for projects and organizations to go to for consulting and development support for a Web site to meet their specific needs, while maintaining a certain level of integration with the enterprise JPL Web space. Customized Web sites, or *portals,* could be tailored to provide subscription services, news headlines, real-time query connections to databases, dynamic or managed links to information resources, and institutional search services. Eventually, user-defined agents would be available for scouring the intranet and internet for timely information feeds related to user-defined subject interests. Typical functions the service would provide include

- Web graphic design and consulting services, including standards and guidelines for Web site designs and restrictions on graphics and logo use
- Web page metatag standards development and publication. Use of metatags in HTML pages greatly enhances relevancy of automated search and retrieval mechanisms. Coordination of metatag syntax with core document and data management metadata standards would provide additional integrated search improvements.
- Web design guidelines. Provide recommendations for designing sites, including navigation tips, HTML versions, multiple browser support, legal requirements.
- Web site publishing tools. Provide templates and procedures for customers to build and manage their own Web sites. Evaluate and recommend procurement of automated Web site management tools to simplify and speed development and maintenance of Web sites.
- Dynamic Web publishing. Provide support for developing Web sites linked to dynamic database content. Provide development tools, publishing access protocols, interface standards, and example code sets.
- Provide guidelines and procedures for Web site configuration management, maintenance, and long-term site archival
- Provide collection of meaningful metrics and generation of customer-requested reports for site use and analysis
- Provide user or group-customizable Web views into JPL internal and external knowledge resources

Category—Infrastructure Applications Service

Requirements—The Web Site Management Service shall

- Develop and publish standards for JPL Web site design and development
- Provide a standard way to link Web sites to databases for dynamic content publishing
- Provide Web publishing capabilities for public, partner, enterprise, or workgroup (e.g. Project, task) dissemination
- Satisfy ISO 9001 compliance for Element 5, Document and Data Control
- Provide standard metatag definitions and procedures for including metadata in Web pages
- Embed access rights labeling in the publishing process
- Provide an API for server-to-server Web page transfer
- Utilize EIS unique user names and principles
- Integrate with the EIS authentication and authorization services, where possible
- Interoperate with EIS messaging service for publication notification, where possible
- Utilize the EIS File services, where possible
- Utilize the EIS Data Access services for database support
- Utilize the knowledge management Identification Service process
- Provide tools for content validation and link verification
- Provide periodic statistical traffic analyses and usage reports based on server logs.
- Provide standard templates for building a custom user or group Web site on the JPLIntranet

Specifications and Standards—HTML, XML and associated standards, Java

Technologies and Products—Photoshop, BBedit, Netscape Composer, Cold Fusion, NetObjects Fusion, Adobe Sitemill, Java, JavaScript, INSO, Adobe PDF, Netscape Netcenter Affinity Portal, Plumtree server, Inso Dynabase, Verity family of products (including Agent server and Profiler), Fulcrum Knowledge Network, Netscape Compass server.

Recommendation—This service should standardize the tools used to build and maintain Web sites by selecting standard tools such as HTML 3.2-4.0 (moving to XML for intranet and extranet use) and HTML 3.2 for public Web sites, with PDF, GIF, JPEG. This service may be a logical “value-added” application extension to EIS.

Rationale—Almost every organization and project at JPL is developing and maintaining one or more Web sites. There are over 150,000 Web pages at JPL. Many are managed by academic part-time students, administrative staff, or engineering staff that are not trained in proper Web site development and management. Sites often go up and never come down, or content becomes stale. Links become broken and are not maintained. Site and data security requirements are often misunderstood or misconfigured. The usability of sites published for public access is often not tested from slow links or through various Internet Service Providers (ISPs) or browser types.

4.5.3 Interchange and Conversion

Description—This service provides an important data translation service for conversion from proprietary to standard, vendor-neutral, or exchange formats. This would include functions to facilitate data import and export between design and development tools and data management systems as provided by the Document and Data Management Service. Data porting and mapping services would be part of this service. Identifies core office software and browser configurations for increased interoperability. Develops and/or recommends naming standards, processes, and standard APIs for facilitating data interoperability and portability. Includes conversion services for data preservation and migration to new media. Typical functions provided would include

- Desktop standards recommendations—Provide an institutionally supported set of desktop and workstation applications and configurations, tested and verified to provide a documented level of interoperability for sharing information. Institutional support would include providing a service base for installation of and help for the desktop standards, and augmenting the standard set as new needs for enterprise knowledge sharing arise (e.g. need to share audio or video files).
- DNS Alliance implementation of standards—Provide service base for testing, acquiring, and installing supported applications and configurations for developers and users
- Naming standards—Provide designed and documented conventions for naming files, Web sites, etc., and user notification assistance for implementations observing these conventions
- Develop and recommend standard APIs—Recommend and (as appropriate) develop published and stable API to selected institutional applications. Purpose is to facilitate implementation and sustaining operation of applications that must communicate with these institutional applications.
- Format and media conversion services—Provide format conversion services for selected combinations of source file/destination file not covered by desktop standards or APIs, such as PDF, PS, STEP, HTML, XML, SGML, RTF, $\TeX$, etc. Provide media conversion services for migration of information from paper or media that is at risk of becoming obsolete or otherwise unreadable

Category—Infrastructure Applications Service[3]

Requirements—The Interchange and Conversion Service shall

[3] Portions of this service were originally proposed as a Data Interchange Service in Reference [__].

- Develop and assist in installation of minimum configurations of core product software on all JPL PC and Macintosh desktops
- Develop and publish standards for UNIX desktop workstations for core product software
- Maintain file access rights in any conversion process
- Provide a set of APIs for data interchange between selected enterprise core applications
- Procure, tailor, and develop conversion tools for native to vendor-neutral file standards, as required
- Provide methods for exchange of documents, electronic designs, graphics, images, and video
- Interoperate with the Document and Data Management Service

Specifications and Standards—HTML, XML, PDF, STEP, GIF, JPEG, Java Enterprise APIs, FlashPix, Internet Imaging Protocol (IIP). In addition, other standards and specifications discussed in the EIS Architecture Specification, Section 3.3.6 should be revisited for consideration. IGES and DXF for example will continue to provide utility for interchange within specific disciplines, and will be part of an evolutionary trend toward STEP.

Technologies and Products—FastTag, Inso conversion filters, Adobe Acrobat Exchange, Verity KeyView HTML Export SDK, Java Development Kit APIs and Products, Live Picture Inc., LEAD Technologies

Recommendation—Review, update, and extend the recommended core product desktop software standards, taking into consideration changes due to the Oracle applications configurations. Support installation of the recommended core set of desktop productivity applications (e.g. Microsoft Office, Web browser, e-mail client, calendar client), versions and configurations to a minimum interoperable level for intranet native file and message exchange. Provide PDF writers on all desktops. Increase use of automated server-based rendering for viewing native formats in open standards such as HTML. Begin work on establishing and automating file conversions to vendor archive formats for translation information that requires long retention periods or is suitable for reuse. ICIS and EIS already perform portions of this function, and these features are a logical extension of this service.

Rationale—It is estimated that JPL workers spend a substantial percent of their time in unproductive activities related to translation and interchange of basic documents. Best practices in industry typically require employees to conform to basic interchange standards, often by maintaining tight control of desktops and the applications they use. At JPL, we will continue to require interoperability between desktop PC, UNIX, and Macintosh workstation platforms and project choice about the use of recommended standards. (Note these are not standards for server-class machines.) Our growing alliances with external partners also necessitate a need for maintaining our dedication to open standards for information exchange. However, it does not make economic sense to perform large-scale conversions on native files that are exchanged frequently during the design and development cycle. When products enter a phase of configuration management and require long-term archive, conversion to vendor-neutral formats becomes a requirement.

4.5.4 Data Archive

Description—The Data Archive Service provides long-term archival and data warehousing of institutional documents and data. It provides periodic and automated checks of active knowledge repositories and resources in order to retire, reduce online material, and archive resources to secondary media and/or storage, such as CD-R or CD-ROM. It will be responsible for maintaining archived records for JPL legal and contractual requirements. The Data Archive Service will develop, maintain, and provide access (electronic and/or physical) to historical archives and records on secondary or hardcopy storage. As part of this process, the service develops and maintains electronic catalogs of all archived records and products for rapid retrieval and location of archived information.

The Data Archive Service also provides support to collect and store quantitative and qualitative data for later analysis, decision support, process improvement, and metrics collection. These functions are sometimes referred to as *data warehousing* capabilities and are typically built on top of robust database technologies, including relational, object-relational, and/or object-based technology. The primary difference between a data warehouse and a general operational database is that the data is organized and built around subjects rather than specific applications. In addition, it does not need to be kept up-to-the-minute as required in a transaction processing system. Data warehouses are useful for analysis with data that may be a day, week, or month old, depending on the analysis required. Different types of data may have different refresh rates. For a data warehouse to be effective, it must ensure the quality of the data within it through data cleansing methods.

A data warehouse can be designed to be as targeted or as comprehensive as necessary. The warehouse can be built one piece at a time, with a focus initially on subjects having the highest priority in meeting current business objectives. A data mart capability can be considered a component of a larger data warehouse and are often subject specific and easier to build. For data marts to be effective, they need to be built with a design geared toward their place within a larger data warehousing architecture.

The logical extension of this Data Archive Service is to move from primarily a business-enhancing support tool to one that includes science data warehousing for science analysis. The vast amounts of science and engineering data that has already been produced by JPL is the basis for a potentially huge science data warehouse. Most of these data are already labeled and classified by subject area and access provided by many distributed data systems.

Category—Infrastructure Application Service

Requirements—The Data Archive Service shall

- Develop and maintain institutional digital document, data, photographic, hardcopy, vellum, videotape, microfiche archives, and electronic catalogs of the archived products
- Provide processes and tools to measure knowledge aging
- Provide services to migrate data from active on-line repositories to archival storage media and/or near-line storage
- Maintain archive media, migrating data as needed to preserve data integrity
- Provide tools for data collection, data porting (imports/exports), data cleansing, data transformation, and data replication/snapshots
- Develop a document and data warehouse architecture that defines a data warehouse data model, metadata, structure, and interfaces
- Provide a consulting, design, and data modeling service for knowledge resource development to facilitate migration to archives after the end of active data lifecycles
- Evaluate, procure, and recommend tools for developing and managing the archives
- Interoperate with the knowledge management Catalog Service
- Integrate with EIS data access service for data storage
- Integrate with EIS authentication and directory services, where possible
- Provide standard mechanisms for data and metadata exchange with other applications and knowledge management services

Specifications and Standards—ISO 9660 for CD-ROM and CD-R, SQL, XML, MDIS, JDBC, ODBC

Technologies and Products—CD-ROM, CD-R, DVD for long-term digital archives. Optical tape, Oracle Data Mart Suite, Oracle Data Warehouse, Visible Advantage, Erwin, Broadbase Information Systems, Sagent Data Mart Solution, Silvon Software Inc. DataTracker, MicroStrategy Inc DSS Architect, Informatica Corp. PowerCenter, Prism Solutions Inc. Prism Warehouse Directory and Prism Web Access, Red Brick Warehouse, Information Builders Inc. SmartMart, Logic Works Universal Directory

Recommendation—Build on existing Data Archive Services and expertise and address the new requirements for JPL compliance with ISO 9001 and NASA NPG 7120.5A. Focus initial effort on updating and creating high-level catalogs that point to non-digital archives. Adopting internal best practices for creating CD-R archives and begin capturing end of project documentation on CD-R prior to transfer to vellum files where they are not available digitally. Review and update policies of the Engineering Documents and Drawings to ensure they meet ISO requirements. Prioritize non-electronic material for digitization, based on metrics collected for frequency of user requests. Utilize in-house expertise of the Data Distribution Laboratory, and science data archive systems (e.g. PDS, PO.DAAC, MIPS) to apply to institutional archives.

For data warehousing, work with NBS to evaluate use of Oracle-based products for business data warehousing needs, engineering, and science large-scale warehousing needs. NBS has already indicated an interest in providing some level of data warehousing for providing business to external applications and users. Evaluate vendor-neutral data warehouse solutions (e.g. RedBrick data warehouse tool suites) for architecting and designing the framework for a Data Archive Service at JPL that incorporates data warehousing. Perform economic analysis of remaining with single vendor (Oracle) versus the cost of interoperability, maintenance and training for multiple technologies.

Rationale—Database support at the enterprise level is expensive and fragmenting administrative personnel by purchasing products and technologies from many vendors is generally not cost effective. The ability to perform management decision support services are predicated by the existence of a large body of information, typically stored in a data warehouse for analysis. Most of the commercial emphasis on data warehousing has been with high volume transaction processing, and the storage of vast amounts of purchasing information that companies can then use to tune and target their products and services to specific customers.

4.5.5 Catalog

Description—The Catalog Service provides consulting services and simple tools to create and maintain catalogs of information or *knowledge resources*. Catalogs are most useful when logical organization structures are used, so this is a service that benefits a great deal from traditional library science methodologies and application of standards. By developing standard categories and simple taxonomies for organization of enterprise information and knowledge, the service plays an important role in the development of a JPL *Yahoo*!-style portal.

Category—Knowledge Management Service

Requirements—The Catalog Service shall

- Develop and publish recommended standards for institutional data categorization
- Develop and publish a minimum core set of metadata standards for all enterprise document and data objects
- Create and maintain a Lab-wide data dictionary
- Develop and maintain an enterprise catalog of information resources and their attributes
- Include an electronic forms capability
- Integrate with the Search, Browse and Index Service
- Provide consulting for new catalog and knowledge base development

Specifications and Standards—XML, MDIS, OIM, Dublin Core, MARC

Technologies and Products—This service could utilize the electronic forms for catalog development

Recommendation—Move toward XML for metadata transfer between systems. Focus on development of a very simple classification framework for enterprise documentation and data, beginning with work already done for JPL and NASA Technical Libraries, DMIE, ELIAS document library coordinators for core metadata requirements for ISO 9001 compliance, without requiring strict adherence to MARC or Dublin core standards. Utilize automatic metadata extraction mechanisms whenever possible. Use of keyword assignments by end users (as opposed to catalog specialists) should be minimized and used as part of a controlled vocabulary or taxonomy.

Rationale—Research has shown that in engineering organizations, a flexible and scaleable approach to metadata assignment and standardization works best. The vendor development efforts focused on open metadata models and simple mechanisms for metadata exchange are particularly promising.

4.6 Knowledge Distribution Services

4.6.1 Identification

Description—The Identification Service provides a standard process for establishing single unique user identity and associated user information and status for all JPL employees, contractors, and partners. This information is required prior to granting access to any knowledge or information resources at JPL. The process ensures that changes in master JPL identity records propagate to the appropriate JPL services and processes. The goal is to provide a common identification capability across many applications and services, including IBS Oracle applications, badge readers, and EIS services (DCE, AFS, NT domain, e-mail, remote login, and Web services). Some of the functions provided would include

- Single user name process
- Single password authentication
- Standard groups
- Standard roles
- Authentication/authorization model and APIs

Category—Knowledge Management Service

Requirements—The Identification Service shall

- Utilize EIS unique user names
- Support all knowledge management and Enterprise applications and services
- Interoperate with a designated EIS security service for authentication

- Interoperate with the EIS directory service for user information
- Interoperate with the IBS Oracle applications services
- Design and publicize standard conventions for creating and naming groups and roles that can be incorporated into applications

Specifications and Standards—NASA Strategy for Windows NT Domain (NASA-STD-2801), NASA Firewall Strategy, Architecture, Standards, and Products (NASA-STD-2813), LDAP, X509 certificates, Secure Sockets Layer (SSL)

Technologies and Products—Virtual Private Networks (VPN), RAZ remote access, X509 certificates, and Entrust PKI.

Recommendation—Work closely with EIS and the JPL Security group to make progress toward the original EIS goal of single sign-on and the NASA plans for establishing an agency wide public key infrastructure. This may be achieved by a level of synchronization between multiple security servers, each designed for servicing a subset of JPL applications. Utilize IBS systems as “gold” sources of information about badge-holding JPL and contractor users. Utilize EIS Directory service as the central repository for up-to-date user information available via an LDAP interface to any application. Where possible, integrate applications to one of several synchronized EIS security services for user authentication.

Rationale—There has been an increasing number of user and group account maintenance requirements as the move to more COTS solutions are deployed without consistent APIs to link to a common security framework. As a result, most users have several passwords for accessing different applications, creating a burden for users and potential security concerns.

4.6.2 Search, Browse, and Index

Description—This service provides enhanced Web-based mechanisms to locate and link both internal and external knowledge resources of interest to an individual to support his or her daily work. The focus will be on providing rich access to internal resources of all types, by utilizing many of the features found on the Web through use of advanced search engines and through logically organized portal sites. A service like this provides both full-text and metadata search and indexing capabilities for both structured and unstructured data found in a variety of resources. It provides a variety of navigation methods to meet preferences of different types of users. The service takes advantage of increasingly powerful technologies that utilize spiders, robots, crawlers, clustering, summarization, and gist-extraction techniques to provide targeted information to users without a heavy burden on labor-intensive tasks such as manually creating key words. Typical functions this service would provide include

- Organized folder navigation for browsing JPL knowledge resources
- Index, categorize, and search JPL Web sites, databases, and file systems
- Index, categorize, and search JPL distributed document and data management systems

- Basic and advanced search capabilities across all or selected knowledge resources
- Canned queries and reports for general information
- Advanced navigation techniques (e.g. hyperbolic tree) and natural language queries

Category—Knowledge Management Service

Requirements—The Search, Browse, and Index Service shall

- Provide a variety of navigation mechanisms for finding knowledge and information from JPL repositories and knowledge bases
- Provide mechanisms for groups and individuals to register knowledge repositories for indexing and access
- Integrate with EIS authentication services for access to knowledge resources
- Interoperate with the Catalog Service
- Be engineered, deployed, and staffed for continuous, high availability operation

Specifications and Standards—None. Primary technologies for indexing and categorization are proprietary. However, XML provides good content representation.

Technologies and Products—Verity Information Server, Fulcrum Knowledge Network, KnowledgeX, IntraSpect, DataChannel ChannelManager

Recommendation—Improve and augment current internal use of the Verity indexing product provided with ELIAS. Extend to provide indexing service for a variety of intranet repositories, including DMS, file systems, databases. Develop categorization classes utilizing Verity customizable topics technologies. Later evaluate Verity extensions for profiling and agent services. Utilize a cross-functional team, including information systems and library science professionals, for developing this service.

Rationale—The Verity product has already been purchased and is being underutilized. This product is also in use at NASA headquarters, and it may be possible to collaborate on information categories that will help gear indexing of public material toward the major NASA themes as well.

4.6.3 Research

Description—The Research Service provides technical librarians for expert reference capabilities for the JPL worker community. Reference and research options include, but are not limited to: information retrieval, information synthesis, information distillation, competitor intelligence, document delivery, and information organization guidance.

Category—Knowledge Management Service

Requirements—The Research Service shall

- Provide expert help at navigating through complex, highly technical externally produced paper and online resources
- Provide mediated manual and online information retrieval tuned to the needs of the Laboratory communities
- Provide knowledge of search mechanics and retrieval theory, using a range of diverse and sophisticated state-of-the-art commercial and government bibliographic and full-text retrieval systems
- Provide applications of library science terminology for foreign language and discipline-oriented vocabularies—including information sciences, physical sciences, life sciences, engineering, management sciences, and law
- Augment existing technical library services with unmediated services
- Analyze needs of science and technical staff during various project and proposal phases for highly technical outside information resources
- Improve ways scientists and engineers include expert navigation of externally produced paper and online resources into proposal, design, and development processes
- Research ways and potential benefit of providing targeted and automated external information feeds directly to users

Technologies and Products— Sirsi WebCat, many others

Specifications and Standards—MARC, Dublin Core, Z39.50

Technologies and Products—Many.

Recommendation—Continue and augment existing research and reference library function. Involve technical staff in other knowledge management services, particularly the Catalog, Connection, and Search, Index, and Browse Service, as well as the Document and Data Management Service.

Rationale—Technical specialists trained in library cataloging and search methodologies should have much to offer by working closely with information systems professionals in the use and deployment of advanced computer technologies for information retrieval.

4.6.4 Information Analysis and Mining

Description—The Information Analysis and Mining Service provides tools, technologies, and methods for processing and interpreting qualitative and quantitative data. This includes query and reporting mechanism for drilling down into large quantities of detailed or summarized data. It can provides statistical, trend, demographic, and theme analyses useful for decision support. An information analysis and mining environment would provide products, services, methodology, and partnerships for decision support that helps users discover valuable knowledge from very large data sets for both business-critical and scientific questions. While most commercial data mining and analysis tools have focused on business data sets, the techniques can be applied to science and engineering data sets as well. Some of the functions provided would include

- Data mining process development
- Identifying relationships between data mining and other scientific disciplines
- Data mining problem formalization
- Data preprocessing, classification, and clustering
- Database structures and their operations
- Time serial data analysis
- Data visualization
- Prediction and forecasting
- Query tools
- Report generation tools
- Quantitative and qualitative data analysis packages
- Online analytical processing (OLAP) software

Category—Knowledge Management Service

Requirements—The Information Analysis and Mining Service shall

- Interoperate with the Data Archive Service
- Develop a data mining and analysis process
- Provide a set of tools and supporting services for data mining and analysis functions.
- Interoperate with the EIS security service

Specifications and Standards—SQL, SQL3

Technologies and Products—Oracle Discoverer, Red Brick DecisionScape, Red Brick Data Mine

Recommendation—Start by providing simple tools for query (drill down), time series analysis, and data visualization that can apply to both science and business data sets. For science, engineering, and image analysis and data mining, make use of expertise in Division 36 (previously 39) for advanced data mining and analysis techniques (e.g. pattern recognition) that could apply to institutional data sets.

Rationale—The greatest benefit of having knowledge available electronically is being able to find, filter, and quickly analyze that knowledge for those things that are pertinent to the user. Information analysis and mining tools and techniques will provide powerful capabilities in those areas.

4.6.5 Workflow

Description—The Workflow Service provides general electronic workflow and electronic forms capabilities for automating routine processes. This includes the ability to set up real-time notification parameters, reminders, and delegation services. Some of the functions the Workflow Service may provide include

- Wizards for guiding a user through entering information through a series of easily answered questions. These often include real-time validation of the data entered.
- Tools for setting up routine work flows to match predictable work processes. This would include the definition of notifications (e.g. e-mail) to specific users or groups, and the ability to delegate action to others.
- Publish and subscribe capabilities for both end-users and groups

More advanced workflow or automation capabilities would apply to ad hoc or dynamic work processes. These are more closely related to artificial intelligence and expert systems categories, and would be considered long-term or future directions for a knowledge management workflow (or automation) service.

Category—Knowledge Management Service

Requirements—The Workflow Service shall

- Provide an electronic approval and authorization capability
- Provide an open API for integration with existing applications
- Provide a tool to create custom workflow
- Provide a capability to link nodes in workflows to user or group recipients
- Provide tools to create data collection and validation “wizards”

- Interoperate with the Web Site Management Service for Web-based data entry capability
- Provide a design mechanism for creating user-defined forms
- Provide a workflow notification component to facilitate standard routing
- Provide storage for form templates and completed forms
- Integrate with the EIS messaging service for notification
- Interoperate with the EIS directory service for user information
- Integrate with EIS data access service for database storage
- Integrate with EIS authentication and directory services, where possible

Specifications and Standards—WFMC, object models, and NASA-STD-2809 Intelligent Electronic Forms

Technologies and Products—Many COTS products have workflow built into them for a specific purpose. Examples include electronic forms routing and publishing features in Document Management systems such as OpenText Livelink, Help Desk products such as Remedy ARS, and paging products. Workflow related products that provide a messaging transport at a lower level that enable workflow include Tibco and MQ Series. Electronic forms products include FormFlow filler, JetForm, Shana Informed, Caere OmniForm, and Livelink with integrated JetForm.

Recommendation—Review JPL's current use of several forms management systems and servers and the recommended NASA-STD-2809. Select and manage a single enterprise standard system. Include an evaluation of existing forms capabilities that may have been introduced with the new Oracle applications suite, since many of the current electronic forms are business forms. This evaluation should help identify where a likely home and service base for the application should be with EIS, IBS, or Division 64. The JetForm product architecture and direction appears to be consistent with the knowledge management open-standards approach. Continue to utilize existing workflow components in existing applications, working to integrate these with EIS services for common user and group information.

Rationale—A single Web-based forms management system together with a service base for easily designing new forms would be useful to a variety of productivity improvements. The electronic forms service already exists in Division 64 and utilizes the FormFlow product, originally purchased from Delrina. This product is now part of the JetForm product line; Oracle is one of JetForm's key partners. Over 100 forms are available online, and Windows and Macintosh clients are available for online data entry. Human Resources also used a JetForm server for authorizing changes to timecard data, before the switch to the new Web-based timekeeping system. The current Forms application does not appear to be well known or used outside of the administrative area, but there is no reason it cannot be used effectively for management and engineering purposes (e.g. engineering change requests, problem failure reports, action items, or configuration change board activities). Forms should be integrated with a database for storage and retrieval. There has been a consolidation within the industry over the past few years, and many products now provide Web-based forms.

General workflow standardization is another issue. The industry for stand-alone workflow products and standards is immature. The WFMC standard has been shown to work between a few pilot vendors, but the complexity of the standard has made its adoption slow. Workflow interoperability may be obtainable in other ways in the future, including the use of XML and simple APIs such as SAX for event-driven data exchange.

4.7 Knowledge Management Services

The Knowledge Management process is recommended as a core process in the Provide Enabling Services (PES) domain. Part of the responsibility of the knowledge management process owner will be to understand the environment that knowledge management must be responsive to at JPL. This includes detailed knowledge of the legal and contractual obligations that a knowledge management system and service base must incorporate. For this reason, one of the primary knowledge-management related services is assigned to the general knowledge management process.

The following knowledge management services provide underlying capabilities that all of the knowledge management services need and would utilize.

4.7.1 Training

Description—The Training service provides user training, communications, and user groups for effective use and creation of institutional knowledge management resources. The functions provided by this service include

- Training in contributing to and using JPL knowledge resources
- Communications, including newsletters, announcements, and articles on the knowledge management service availability and resources that have been added to knowledge resources

Category—Knowledge Management Service

Requirements—The Training Service shall

- Provide basic and advanced instruction on the creation and use of the JPL knowledge resources
- Maintain open lines of communication within the JPL community, informing them of upcoming knowledge management improvements and soliciting feedback to improve the knowledge management and applications infrastructure services

Specifications and Standards—None

Technologies and Products—None

Recommendation—Utilize existing communications and training mechanisms, including Professional Development, ICIS News Bytes, e-mail, inclusion of knowledge management resources and training in orientation materials, a knowledge management web site and document library, and regularly scheduled noon-time talks. Pursue offering executive level short-course on knowledge management from outside consultants.

Rationale—Continues philosophy of utilizing existing services

4.7.2 Operations and Maintenance

Description—The Operations and Maintenance service provides hardware, software, system administration, help desk services for processes, projects, and groups utilizing knowledge management and/or infrastructure application services. Assists users in desktop configurations and installations. Updates existing knowledge bases with additional metadata, keywords, and cross-references as part of standard operating procedures. Routinely collects and reports use metrics for knowledge repositories.

Category—Infrastructure Applications Service.

Requirements—The Operations and Maintenance Service shall be

- Engineered, deployed, and staffed for 24 by 7 continuous, high availability operations.
- Staffed for standard 8 by 5 work weeks

Specifications and Standards—None

Technologies and Products—Tivoli TME 10, paging services, Remedy ARS, DNS Alliance

Recommendation—Utilize EIS and DNS Alliance operations staff for in-house developed services. Services selected for partner development may be operated and maintained by the partner.

Rationale—The knowledge management and infrastructure application services are designed to build on top of existing infrastructure services. Help Desk, systems administration, and operations and routine procedures are already part of the EIS and DNS Alliance charters for enterprise level support.

4.8 Overall Knowledge Management Process

Overseeing the knowledge management process requires the process owner to assess constraints, interpret and clarify rules, ensure that the services continue to be provided and integrated, and to verify that incentives for knowledge contribution, use, and reuse are encouraged. While the knowledge management process is not a service, it is mentioned here to capture those functions that the process owner will need to perform, such as

- Collect and review all rules, regulations, and policies related to how JPL manages its knowledge
- Interface with the Office of the General Council, JPL Security, Contract Management Office, Acquisition, Technology Transfer and Utilization Office, Software Dissemination office, and other offices to obtain and interpret and clarify specific requirements
- Interpret the impact of those rules on knowledge management policies for data and document classification and access
- Publish derivative policies aggregating, synthesizing, and clarifying rules to match JPL work processes
- Assist process owners in updating their procedures and policies as required by new and/or changing policies

Some of this work is already being done, but additional coordination, integration, and dissemination of important information is required. It is difficult for projects and employees to understand what their specific responsibilities are when it comes to publication and sharing of data, information, software, and designs as part of their daily work. Regulations on ITAR materials, public release procedures, intellectual property rights, copyright rules, and computer security are just a few of the items that need to be clarified for JPL employees and affiliates.

The process owner will need to

- Develop and maintain a master list of rules, regulations, and constraints applicable to JPL knowledge, information, and data management
- Maintain a list of JPL or Caltech “subject matter experts” who are responsible for interpretation of the rules
- Clarify and publish JPL policies for the proper creation and management of JPL knowledge resources
- Continue and improve coordination between groups, including the Electronic Communications Committee, JPL Computer Security, Network Operations, Media Relations Office, and the Office of the General Council

5 Information System Architecture

The information system architecture presents a generalized methodology, a standard approach, to system architecture development and implementation for the JPL Knowledge Base. This approach, although it may not utilize the latest technologies to optimize some localized solutions for network computing, nevertheless seeks to provide a structure that optimizes the knowledge management services at the enterprise level, and applies to all knowledge resources in JPL's Knowledge Base (i.e., it does not apply to specialized, local information services).

In this sense, this system architecture is a general purpose architecture that can satisfy, at best, only a majority of users (not all). These users may be JPL employees or contractors, partners, subcontractors or vendors, and may be US citizens or foreign nationals. Conversely, there are also people who must be denied access to this information[4], for many, different reasons. Therefore, this ISA identifies standardized methods for providing information to authorized users in a secure environment.

This information system architecture adopts a modular approach to system development where certain functions can be assembled into components that may be reused throughout the entire system. An example would be a security module that is invoked when a user attempts to access a sensitive document. This module may need to contact two or more servers in order to authenticate such a user as well as to verify the access privileges associated with that specific document. Additionally, the document itself requires a standard set of metadata and mechanisms for accessing the metadata that allow other applications to assess its security and access privileges.

Several people will "touch" information as value is added to it in the course of its lifecycle, from creation to eventual archival and, perhaps, disposal. For instance, in the case of a document being published, this touching may come about through several events: authoring, editing, approving, releasing, adding information about the document (metadata) to a database, copying, revising, printing, and transmitting. These people may access this document from any number of applications and/or platforms. The security and authentication mechanisms through which these events take place should not be re-created for each of these applications or platforms; rather they should be uniform and consistent throughout the enterprise, thus ensuring the accuracy and fidelity of information no matter how it is used.

[4] The word "information" as used in this section refers to any data, document, record, image, or file in hardcopy or electronic form that resides in the JPL Knowledge Base.

5.1 System Architecture

The system architecture utilizes a layered structure: user interface, knowledge management functions, application infrastructure services, and knowledge resources, operating on top of existing Infrastructure Services (Figure 5-1).

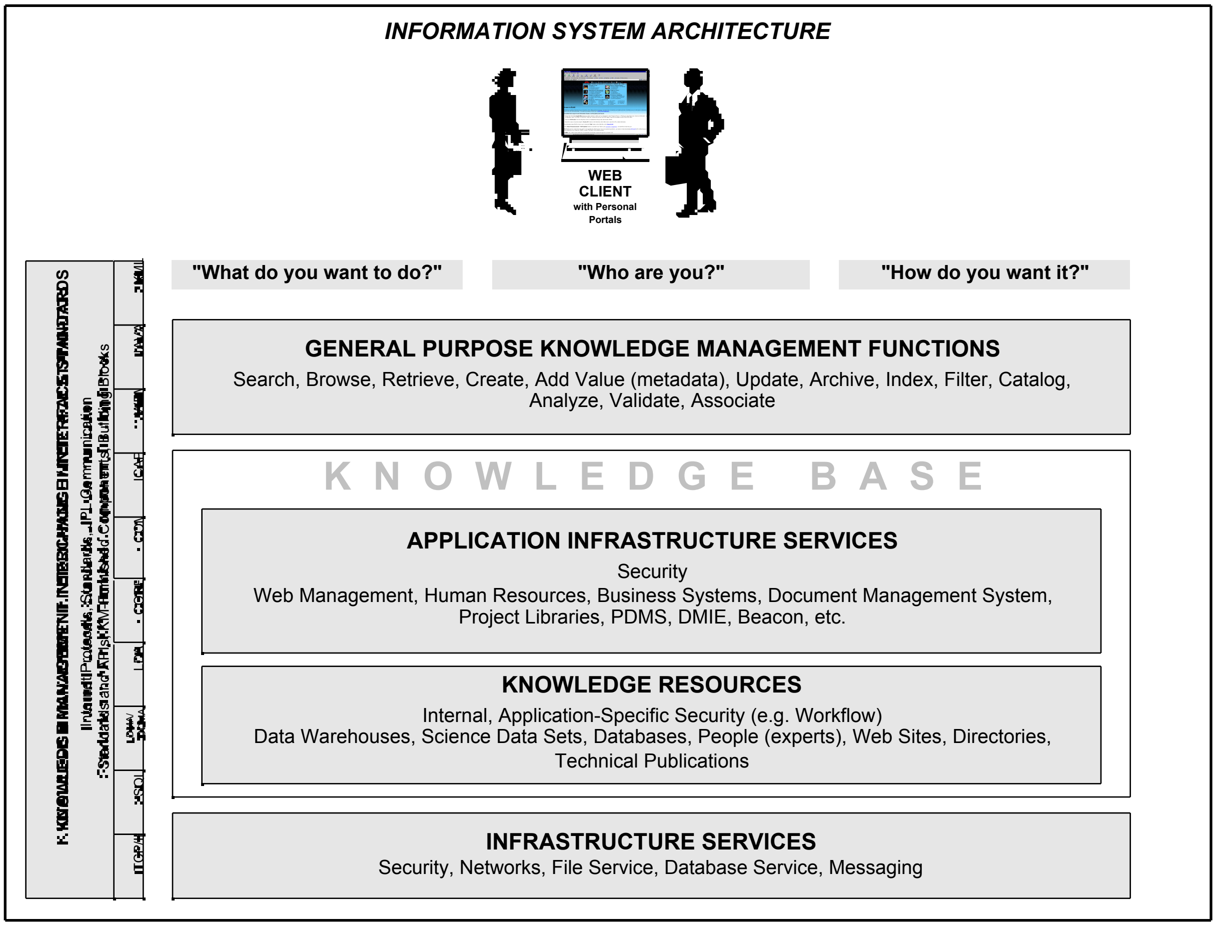


Figure 5-1. High-Level System Architecture

5.1.1 User Interface

The user interface (UI) is the point at which the user interacts with the knowledge management system. This UI is Web-based, meaning it utilizes standard Web browsers (e.g., Netscape Navigator and Microsoft Internet Explorer) as the Web client. The purpose of standardizing on a Web browser for the user interface is two-fold: it reduces overall maintenance costs (deployment and upgrades), and it provides a standard look and feel. The interactive features that use the same mouse and keyboard actions, as well as screen functions, such as scroll bars, hyperlinks, and drop-down lists are the same from one Web page to another. The familiarity of user interface functions that are shared among Web-based applications reduces the learning curve and, hence, training costs. More and more, the browser will use Java applets as they are downloaded from Web servers for application specific functions, and these applets will utilize the same user-interface functions as Web pages do now.

The default user interface should be customizable so as to allow each user to tailor his or her own interface to suit one's particular needs. The phrase "personal portal" has come to mean the user's own home page, personally customized to allow quick access to frequently used Web pages, documents, or other services provided by the knowledge management system.

5.1.2 Knowledge Management Functions

The general-purpose knowledge management functions are common to most Web-based applications. The system architecture defines standard functions that may be encapsulated in components or APIs that are developed as part of the interface specification for each Knowledge Resource. These functions are actions (verbs) that a user invokes and/or applies to information in the knowledge management system. Some of these functions are search, browse, retrieve, create, add value, update, archive, index, filter, catalog, analyze, validate, and associate.

Standard methodologies and mechanisms for implementing these functions using standard naming conventions and published specifications will be maintained in a knowledge management database available to all authorized users (programmers and applications) through the use of standard protocols.

5.1.3 Application Infrastructure Services

The application infrastructure services are defined as a variety of standard and Web-server applications that provide connections between the user interface (i.e., a Web browser) and the underlying knowledge resources. Since many of the functions (described above) require queries to more than one knowledge resource, it is desirable to have standard query functions encapsulated in components that can be reused easily.

Implicit in this design is the use of application service directories that provide reference to physical locations of specific functions (such as security and authentication) or other information when they reside on servers other than the local machine. This will enable various services to be distributed throughout the enterprise, while eliminating the need for specific addresses to be maintained on all the applications needing such services. For example, this would enable the retrieval of a document by reference only to its document number, regardless of the server on which the document actually resides.

The Web servers in the knowledge management system will utilize standardized Web services for the functions described above. All Web servers in the JPL Knowledge Base will adhere to the specifications supported and documented by the knowledge management Operations and Maintenance Service. This will reduce the overall cost of deployment and maintenance and will enable JPL to leverage its investment in server administration and expertise, (i.e., system administration). Standard server applications, servlets, and protocols will provide an environment that promotes reuse of software.

Some examples of application infrastructure services are: Web site management, human resources, business systems, document management systems, project libraries, PDMS, DMIE, and Beacon (the JPL Library).

5.1.4 Knowledge Resources

The Knowledge Resource layer may be considered the data layer, the layer where the actual information resides. Some metadata and indices may reside in a Web server or in some full-text engine. Commercial-off-the-shelf (COTS) products such as document management systems may maintain their own indices and metadata. but have a primary function to manage the actual information objects, such as files, models, and images. Nevertheless, to interoperate in a federated system, there must be communication services (APIs) available to allow queries and responses to and from each knowledge resource.

Other JPL-specific knowledge resources may be developed wholly in-house. For example, a subject-matter expertise (SME) database would be used to locate plasma-physics scientists. This database may have schema and business rules designed specifically for this application; however, it would still comply with security requirements and naming conventions as implemented on other JPL knowledge resources, and would allow for queries posed in a standard format.

Some examples of knowledge resources are: data warehouses, science data sets, databases, people (e.g., subject matter experts), Web sites, directories, product definition data, and technical publications. Infrastructure services shall be provided by EIS. These services currently consist of Security, File, Messaging, and Data Access (Oracle systems).

5.1.5 Knowledge Management Interchange/Interface Standards

Knowledge management interchange and interface standards are used in the development of the components that make up the building blocks within the information systems architecture. A common use of this concept is in application specific interfaces or APIs. These provide standard communication protocols for networked applications that remove the need for comprehension of the underlying applications. Queries and responses are made in standard formats and adhere to standard naming conventions.

The knowledge management system will leverage the world-wide acceptance of internet protocols for use internally at JPL and for communications externally. The knowledge management system will also develop and provide components that allow applications to communicate in a federated system. An example of such a component is the security and authentication module that will be utilized on all Web sites for the purposes of identification of users and verifying their access privileges.

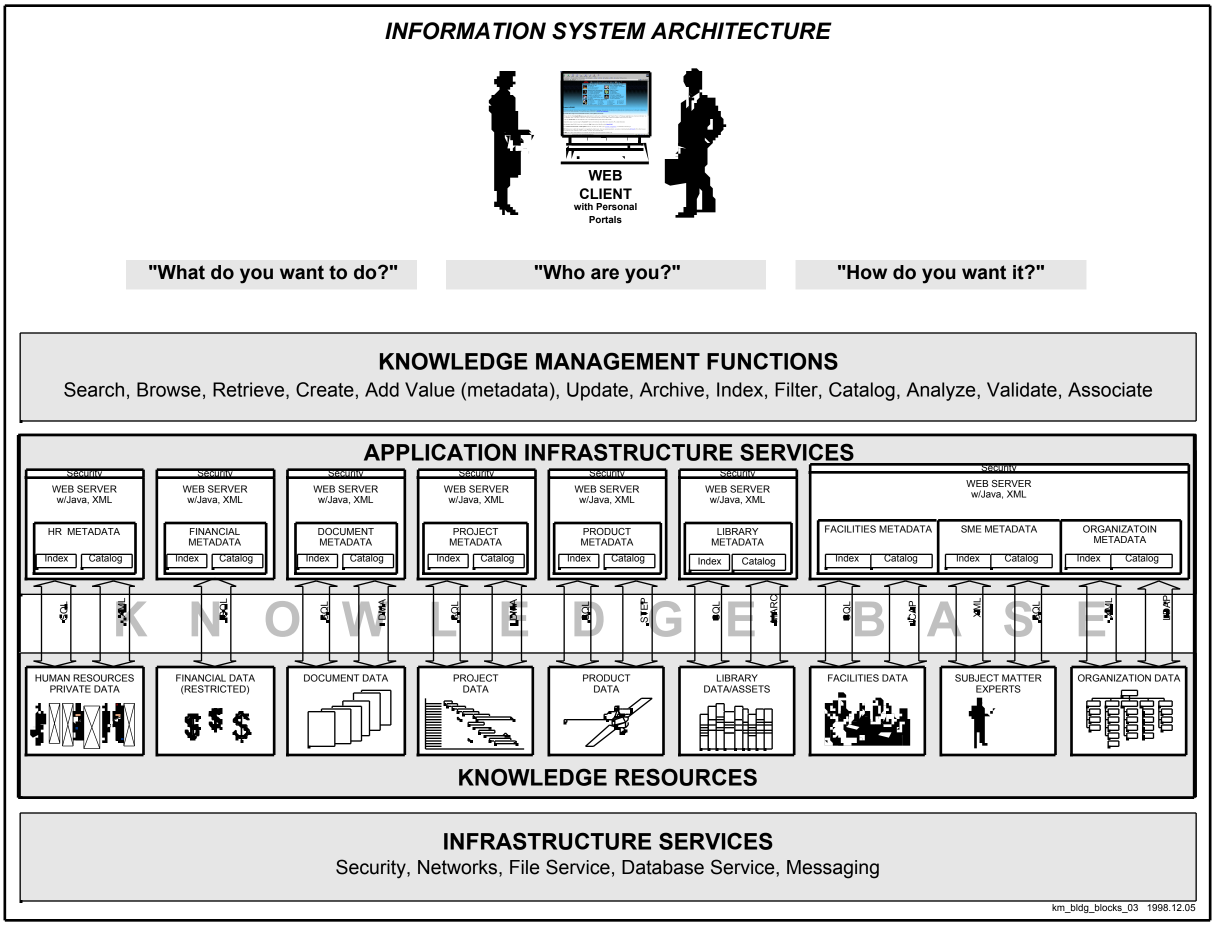


Figure 5-2. Functional System Architecture

Some interchange and interface standards are: TCP/IP, SQL, DMA/ODMA, LDAP, CORBA, COM, ICAP, WFMC, and MARC. These standards are described in greater detail in Appendix B. Additionally, the knowledge management system will standardize on certain development methodologies; specifically the use of the Unified Modeling Language (UML). Additionally, for Web services, the use of Java as a standard programming language will be supported. While other programming languages such as C++ and Perl have been successfully implemented in Web servers, the requirement for software reuse in an enterprise-wide Web-centric architecture overrides choices based on personal preferences. The investment made in standardized development practices will greatly benefit JPL in the long run as systems begin to seamlessly interoperate. Development and maintenance costs should be reduced as a result of this requirement.

Together these layers provide the components necessary for a system architecture. A more detailed view of a functional system architecture is shown in Figure 5-2 that shows the relationship between the distributed knowledge resources and the centralized functions.

5.2 Data Architecture

If the knowledge management system provides the highways upon which the traffic can flow, the data architecture provides the guidelines regarding the types of cars that can travel upon the highways. A key to sharing information across the enterprise is having that information usable and understandable to the person who receives it. Just knowing where it is (e.g., on which server or at which URL) is not enough. The user also needs to be able to read it, save it, manipulate it, and send a new version in a reusable format back to the repository when appropriate. To accomplish these tasks, the data within the Knowledge Base needs to follow certain requirements and standards. The following sections help to identify, within the knowledge architecture:

- Where and by whom data is maintained
- The standard formats data should be kept in for easy reuse
- How data should be labeled for cross-platform compatibility
- How data can be accessed through the knowledge architecture
- In what manner data should be exchanged between knowledge resources
- How data can be described consistently across the Knowledge Base

The data architecture, in conjunction with the system architecture, needs to describe the logical relationship of JPL's data to JPL's business processes [12]—showing the linkage between JPL's goals and objectives and the applications, databases, and rules that flow from those goals and capture JPL's working knowledge. The importance of using a framework (such as the Zachman framework [12]) is the explicit recognition that JPL's goals and objectives should be driving development and implementation of specific systems with specific rules.

5.2.1 Data Sources and Owners

The sources of data at JPL are widely distributed and are operated by and for line, project, and programmatic organizations (Figure 5-3). To develop an architecture that spans these and other external sources, it is critical to understand the scope, nature, and issues related to JPL-specific data storage and ownership. Section 7 describes the types of knowledge resources that exist at JPL; this section looks at how those specific resources are being made available electronically. One key to the knowledge management architecture is that while the knowledge resources can remain in the hands of individual projects, at some point those resources become part of JPL's Knowledge Base—becoming a type of distributed data warehouse comprising all metadata and some information objects. In order to become part of the Knowledge Base, repositories need to adhere to the knowledge management standards and procedures. As part of the Knowledge Base, repositories can greatly increase the number of their users.

In general, data is created, maintained, and used at a local level on a specific project or task. Exceptions to this are those projects that utilize an institutional service such as AFS shared workspace, PDMS for engineering data, DMIE for policies and procedures, or scientists utilizing one of the Distributed Archive and Analysis Centers. In these cases, maintenance has been moved out of the local level, but often creation and utilization are still localized.

Part of the reason for this localization is that users have access to specific repositories, but not to others, and searching across repositories, with the exception of some of those which have moved to the Web, is problematic. Most projects perceive local control of their data to be "better" in terms of time and, sometimes, cost. Efforts to "institutionalize" these local repositories have repeatedly failed because of these deep-seated perceptions. In most cases, the line or project organizations for which the data is maintained feel they are the "owners" of the data, and are hesitant to share their information with others who have not "paid" for it or to turn control over to a centralized source that may not provide good service to them. The general notion seems foreign to most people that the knowledge we gather as part of our paid positions at JPL belongs to JPL. This cultural barrier to sharing knowledge is critical—any system or data architecture is doomed to fail if projects, lines, and individuals don't perceive that JPL-sponsored knowledge belongs to the institution. One solution is to leave control of the data in the hands of the projects, but to have the projects publish the data in standard ways so that others can access it.

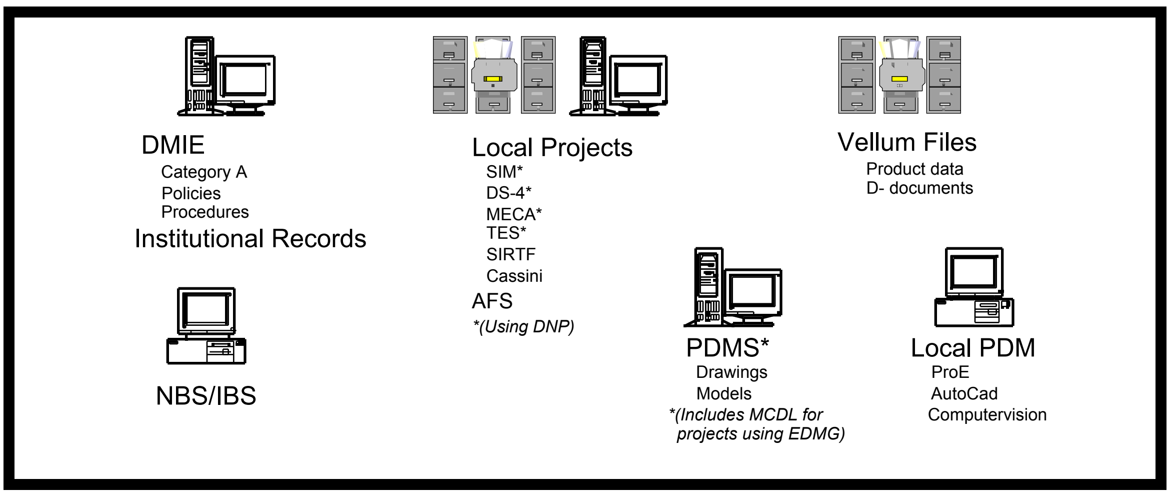


Figure 5-3. JPL Has Distributed Repositories, Some of Which Host Duplicative Data and Are Not Interoperable

5.2.2 Distributed Repositories

An Information Resource Catalog is being compiled by the Enterprise Data Architecture Team, interviewing the owners and maintainers of over 100 information resources across the Lab (see Appendix H). This survey shows some interesting facts

- Most repositories at JPL do not interface with other repositories (except through Web-enabled searches)
- Very little data flows from one repository to another

- Security, configuration management, data archive, and system backup are generally reinvented for each repository, with differing levels of expertise and reference to the business rules governing each area
- Most repositories have a Web-based user interface to the database
- In most cases, a single person maintains the repository

What this implies, in general, is that users cannot search across multiple repositories, and that different repositories will define terms and values in different ways, making it difficult to get similar data from different parts of the Lab. Finally, time and money can be saved by creating building blocks (procedures, software, and/or hardware) incorporating and embedding the appropriate rules that system developers could utilize.

From a knowledge management view, whether the repository is locally controlled or institutionally controlled is irrelevant, as long as certain standards are followed. Therefore, it will be expedient to encourage projects to find a data solution that works for them, minimizes costs to the project and the institution, and allows knowledge to be shared across JPL's interests. It may be determined that in minimizing cost to the institution, some locally controlled repositories should be combined. Certainly, knowledge management requires that a project database or "library" template be prepared with standard building blocks in order to allow projects to quickly meet the diverse requirements levied upon them by knowledge management and other processes.

5.2.3 Standard Data Formats

It is useful to make distinctions among metadata, information objects, and a browsable files

- **Metadata**—Information about the object (e.g., the author, creation date, file format, project)
- **Information object**—The item being described, which may be a reference to a person, an electronic program or file (sometimes called the *native* file, e.g., a detailed model of a star tracker developed in Pro-E), or a pointer to a hard-copy media (video, photograph, or document)
- **Browsable file**—A version of the information object that can be viewed from a Web browser (e.g., the Pro-E version saved in Acrobat of the star tracker model)

To promote ease of sharing existing knowledge, users need access to both a browsable version of an object, as well as to the original version. (A third version may also be required for long-term archive.) The browsable version allows anyone with a Web browser to at least read and see the information in the document, drawing, or object (for example, seeing the values in a spreadsheet). The original (or native) version allows a user with the right software the ability to retrieve and use the native format for full functionality (for example, being able to see the mathematical functions embedded in an Excel spreadsheet).

From a knowledge management point of view, the suite of core-supported software available for desktop workstations should be broad enough to include the majority of functions performed at JPL (such as word processing, spreadsheets, drawings, presentations, and modeling). In addition, file exchange protocols should be published so that users can easily import and export files from those programs.

Basic file formats that can be viewed across JPL's UNIX, PC, and Macintosh platforms includes: HTML, PDF, XML, SGML, and ASCII for text formats, GIF, JPEG, TIFF, and PDF for graphics formats; and MPEG for video.

5.2.4 Interchange Data Formats

The ability for a user to bring information created in one application into another and to reutilize it is crucial for knowledge to be fully reused. To do so requires knowledge creators and knowledge users to have standard formats and interchanges to enable their applications to import and export usable information. Standards exist such as the Federal Geographic Data Committee for geographic information systems, the Dublin Core for digital libraries, and the Common Data Interchange Format within industry. In order to fully share knowledge, data formats should include information that is based on text, photographs, diagrams, numeric, software code, tabular, and video, as well as combinations of those.

As preferred data formats change over time, it is critical that the knowledge management system maintain the ability to read and reuse previously recommended data formats. In general, the trend has been that those formats that rise to precedence do so because they import most standard formats. However, as each iteration of the recommended standards is created, an archive scheme for older formats needs to be maintained.

5.2.5 Data Access

Issues relating to data access derive from both the business rules (who can access what data), as well as the communications, hardware, and software components of the systems involved. Most of the business rule security issues relate to either the user or the data as shown in Table 5-1—these are just samples of the many attributes that will need to be defined for each type of employee access and data type.

Table 5-1. User and Data Access Restriction Issues Derived From Business Rules

User-Related
JPL Employee
—Personnel Title
—Project Title
Foreign National
Contractor with Company A
Contractor with Company B
Principal Investigator on Project A
Data-Related
JPL Discreet

Project D
Task X
ITAR/EAR
Intellectual Property
Proprietary to Company A
Cleared for External Release

The issues relating to the system's abilities involve the data format and interchange issues previously discussed, as well as the access for delivering or retrieving data from that system via a published API. Even if the system itself is using standard tools and formats, without a communication protocol, knowledge cannot be easily shared from that system to the greater JPL community.

5.2.6 Naming Conventions

In adopting basic internet standards, some file naming conventions have already been universally accepted. The three character extensions associated with file names, known as MIME types, enable much of the functionality associated with the Web. The MIME type identifies how the file has been created and how it can be viewed. When organizations are including metadata in their files or are a part of a larger document management system, MIME types are adequate. In the absence of metadata, it is recommended that more structured file naming conventions be followed and that a systematic naming structure be added to the left of the decimal point. This file name would identify the who, what, and when attributes of a file. Some examples of non-structured and structured names are offered below:

Non-Structured Name	Structured Name
pcpa.gif	IMAGE_Thermal-analysis-PCPA.gif
fnamefaq.htm	FAQ_File-Naming-Convention_000000.htm
ha-rates.ppt	SLIDES_New-Holding-Acct-Rates-000000.ppt
toolcat.xls	LIST_DNP-Strategic-Tools-000000.xls
DA-conf-NY.doc	SCHEDULE_NY-Data-Arch-Conf-000000.doc
acctstruc.doc	IOM_GAR-Acct-Struct-000000.doc

Standard naming conventions are needed both within tables and databases, and across the enterprise. By using a consistent naming schema for information objects, both users and system developers will find it much easier to access information over time and to create reusable information. Naming conventions should apply to metadata, data files (both root file names and extensions), MIME types, record types, compound documents, and Web site URLs. Several teams have recommended specific conventions within JPL and NASA. A *single* convention should be agreed upon and recommended across JPL and its partners to the greatest extent possible.

Current constraints have required naming files in the ISO 9660 Level 1 file-naming convention, the 8.3 form *filename.ext*, where *filename* is between 1 and 8 alphanumeric characters and *ext* is between 0 and 3 alphanumeric characters. This works for DOS, Windows 3.X, Windows for Work Groups, Win95, WinNT, Mac OS, UNIX, and VMS. As this limitation is required for compatibility to DOS, Windows 3.X, and Windows for Work Groups, it is recommended that it be superseded by a file naming convention that allows 31 characters in a 27.3 form. This would allow more useful file names, standard extensions, and continued readability with Macintosh (the limiting file system for 31 character file names).

5.2.7 Metadata Standards

JPL's knowledge management architecture is designed to leverage existing object-oriented data structures that include metadata information. Some tools, such as Microsoft Office or Framemaker, and tools that use rich, vendor neutral formats, such as XML, HTML or STEP, encapsulate relevant metadata in the object's data structure.[5] The ability for people to search across JPL's Knowledge Base relies upon the use of consistent metadata collected from each knowledge resource (database). Therefore, there needs to be a core set of metadata elements or attributes in the case of relational databases. The knowledge architecture relies on these standards being in place and is necessary for knowledge management to succeed.

For example, if the element name is "Date", a valid value would be the appropriate date describing the information entered in the proper format. Individual systems could have additional metadata, as long as they maintained the core set with the defined formats. If core software already has embedded tools for metadata, then knowledge management's requirement would be a translator that takes that metadata and brings it into compliance with the Knowledge Base metadata requirements. A basic set of core metadata should require, at a minimum, those attributes noted in Table 5-2.

Table 5-2. Minimal Core Metadata Attributes

Attribute	Description
Title	Title of the information object
Information ID	Alphanumeric identifier (reference designator)
Date	Date of the latest version, revision, or change (YYYY-MM-DD)
Version ID	Alphanumeric identifier of the version, revision, or change
Author	Information preparer(s)
Custodian	Organization or individual responsible for maintaining the information content
Description	Description/Abstract of the information content
Key words	Terms used to catalogue and index the information content

[5] Note that if metadata is stored separately from the data set, changes to the metadata should be reflected in the original data set.

5.2.8 Data Replication and Transfer

One common problem in maintaining databases at JPL is maintaining the currency and accuracy of information within the database. Information that is created and revised locally is generally up to date; however, the information whose parent database is resident in another system is often out of date. Reasons for this asynchronicity revolve around the difficulty in getting access protocols for other systems, as well as maintaining data feeds over time. For example, even after the deployment of NBS, there continue to exist multiple systems, with variant latencies, for finding basic employee location information.

5.2.9 Data Dictionary

Integrating many of the features referred to above, an enterprise-wide data dictionary will allow standard entities, values, and attributes to be defined across JPL to facilitate search, retrieval, and reutilization of information objects. The data dictionary is the authoritative source for data standards and definitions. It should allow for registering and collecting data with related metadata and provide documentation of the lifecycle events for standard data (such as creation, revision, and archive). A key benefit in creating and maintaining a data dictionary is that it quickly identifies duplicative data stores and identifies the authoritative sources for specific data entities.

A realistic data dictionary would have a necessary, but not exhaustive, list of the entities at JPL that cut across disciplines (such as employee, facility, job title, and roles). It is important to understand that various scientific disciplines define the same terms differently. A JPL data dictionary should not try to address all these variations, but should just point to more appropriate, discipline-specific dictionaries for scientific terms (e.g., phase, absolute, and brightness) that relate to data entities. Caution should be taken in trying to compare values for entities like coordinates or angles across disciplines or even across different tools used within a discipline.

6 Implementation Recommendations

6.1 Implementation Planning Process

The knowledge architecture set forth in this document is set at a high level. The implementation planning process is designed to show how the process, service, and system architecture will support the basic intent of knowledge management—to make information available quickly and easily for people to use productively in their day-to-day work at the Lab. Notes compiled from interviews, reviews, benchmarking, suggestions from the JPL community and consultants, and the derived knowledge management requirements were used to determine where new knowledge management methods would be most beneficial.

The prioritized services recommended in this section are intended to provide a solid foundation for an evolving knowledge management infrastructure at JPL. Candidate tasks for initial components of the JPL knowledge management system were considered that are important not only to knowledge management, but to the Lab's basic business of proposing, designing, building, and operating unique space science missions and returning new knowledge to the science community and the public. The initiatives recommended below are those that have a good chance of success, cannot be done without some outside aid, help employees get their jobs done, and support projects' time-critical needs and new ways of working.

6.2 Methodology For Establishing Priorities

In establishing priorities for the specific tasks within the knowledge management services, the team looked at five factors:

- Priority—how urgently is this service needed for the Lab?
- Difficulty—how hard will it be to successfully complete this task?
- Productivity—how much would this service help increase productivity of JPL employees?
- Cost—how expensive will this service be to provide?
- Precedence—what services are needed *before* this service can be deployed?

Part of this evaluation included looking at what services are currently being provided in these areas by JPL organizations, as well as vendors. Specific tasks within those services were prioritized through the ranking process (Table 6-1). These tasks are connected to and will generally be completed as part of one of the ten pilots described in the following section.

Table 6-1. Knowledge Management Services and Prioritized Tasks

Process	Service	Prioritized Tasks
Capture	Resource Development	· Services for design and creation of knowledge resources · Scribing service for key meetings, interviews, and presentations · Metric development and assessment
Develop	Authoring	· Document, engineering, and software development templates, services, procedures, and standards, including templates for word processing, presentations, drawings, models, and software
	Collaboration	· Meeting support tools · Voice and data conferencing
	Connection	· Subject matter experts' directories · Interest groups/forums · Integrated electronic threaded discussions and newsgroups
Organize	Document and Data Management	· Services for document and data publishing (metadata standards, document control/versioning, and access control) · ISO, NPG 7120.5a, and security compliance
	Web Site Management	· Web metatag standards · Web publishing templates and procedures
	Interchange and Conversion	· DNS implementation of NASA recommended standards · Standard data conversion and exchange tools and processes (e.g. native Office, PDF, PS, $\TeX$, RTF, HTML, XML, STEP)
	Data Archive	· Electronic high level catalog of institutional inactive and archived documents and data (including record collections) · Online electronic repository of most frequently requested inactive and archived documents and data (including record collections) · Process to capture end-of-project documents and data for legal, institutional preservation and future access · Provide ISO, NPG 7120.5a, and security compliance
	Catalog	· Core metadata standards development · JPL data categorization and taxonomies · Knowledge Base data dictionary
Distribute	Identification	· Single user authentication process
	Search, Browse, and Index	· Browse, search, and index JPL knowledge resources and Web sites to match standard JPL categories/taxonomies · Search across metadata from multiple repositories
	Research	· External electronic resource subscriptions · Expert research service
KM		· Interpret and clarify current rules (such as NPG 7120, ISO, and ITAR) · Requests for network access to JPL Intranet · Establish incentives for contributions to and reuse of knowledge
	Training	· Training in use and contributions to Knowledge Base · Communications to stakeholders about Knowledge Base
	Operations and Maintenance	· JPL help desk interface agreement with KM services · Routine metrics collection, security verifications, and virus scans

6.3 Knowledge Management Initiatives

The knowledge management initiatives recommended here are intended to focus primarily on support to projects in various stages of their lifecycles—from proposal through final archive of project data. Some of these tasks focus on ensuring that the knowledge management services, processes, and architecture are aligned with the current trends and requirements that might come out of NASA Headquarters, industry, or academia. The intent of these initial knowledge management activities is to address both the prioritized knowledge management services and functions as well as each of the knowledge management business objectives. Taken together, these initiatives focus on helping people get the knowledge they need to do their jobs—whether that knowledge is conveyed in a conversation, presentation, or in an electronic form, and become the heart of the evolving *knowledge management system* at JPL.

To prove the ability of the Implementation Team and the Lab to succeed at knowledge management, it is critical to work on initiatives that affect JPL's core business and integrate a variety of knowledge management services. One of the clear lessons learned from the benchmarking was that phased implementation was a key to success. The activities identified in Table 6-2 are recommended for early implementation and were chosen both for their importance to the Lab's central work, as well as the expected ability of JPL or vendor service providers to deliver the functionality needed. In addition, they provide opportunities for many cross-organizational partnerships at JPL. This is considered a key to the overall success of knowledge management at JPL. These initiatives are discussed in more detail in the sections that follow, and will be expanded in the forthcoming *A Knowledge Management Implementation Plan for JPL.*

Many of these initiatives are interrelated, such that execution of a later initiative requires the successful completion of one or more earlier initiatives. While each task has a primary knowledge management service as a key lead element, many will build upon and utilize from other tasks. Some of those dependencies are mentioned in the descriptions, and will need to be factored into the delivery schedule for the implementation.

By achieving successes in these early knowledge management initiatives, it is anticipated that JPL employees and partners will begin to recognize the advantages of sharing, finding, and knowledge. Contributions and growth of the Knowledge Base will become a natural part of their daily work habits.

Table 6-2. Recommended Knowledge Management Initiatives (the lead service for each initiative is indicated by a bold ***X****).*

Initiative	Resource Development	Authoring	Collaboration	Connection	Document and Data Management	Web Site Mgmt	Interchange and Conversion	Data Archive	Catalog	Identification	Search, Browse, and Index	Research	Information Analysis and Mining	Workflow	Training	Operations and Maintenance
Project Web Site	x					**X**		x	x	x	x				x	x
Project Authoring and Documentation Kit		**X**	x		x			x	x	x	x				x	
Interoperable Libraries	x				**X**	x	x	x	x	x	x				x	x
Collaborative Environment	x		**X**	x			x		x	x	x	x		x	x	x
Expert Connections	x			**X**		x		x	x		x					
Standards Working Group									**X**		x					
JPL Electronic Archives	x				x			**X**	x		x		x			x
Knowledge Navigation				x					x		**X**					x
Knowledge Creation Studies	**X**		x	x								x	x			
NASA Integrated KM Environment			x		x				x	x	x			**X**	x	x
Concurrent Engineering KM		x			x		**X**		x		x				x	x
Scientist/Research Services		x	x	x					x			**X**	x			

The development of the knowledge management system at JPL will be evolutionary and built on the existing knowledge, information systems, and personnel resources that already exist. Early knowledge management initiatives should be phased in as budget, schedule and resources allow. There is also a need to provide an evolving knowledge management system that addresses the cross-section of JPL knowledge management priorities over the early years. This means that although JPL is focused heavily this year as an institution on meeting ISO 9001 and new NPG 7120.5a requirements, JPL should not focus exclusively on tasks that improve efficiency.

Figure 6-1 shows a suggested timing for emphases on each of the 12 early knowledge management tasks identified. This will provide the framework and evaluation measurements to tune the system development as we learn from each major initiative. Each of these tasks should be considered open-ended on the timeline, since there may be some early work done on several of these. In many cases, these initiatives will also continue as their associated knowledge management processes and services are improved. The implementation timeline shows only a three-year window. What we learn during this period will help direct efforts in the years that follow to make good knowledge management practices a central part of JPL's daily work.

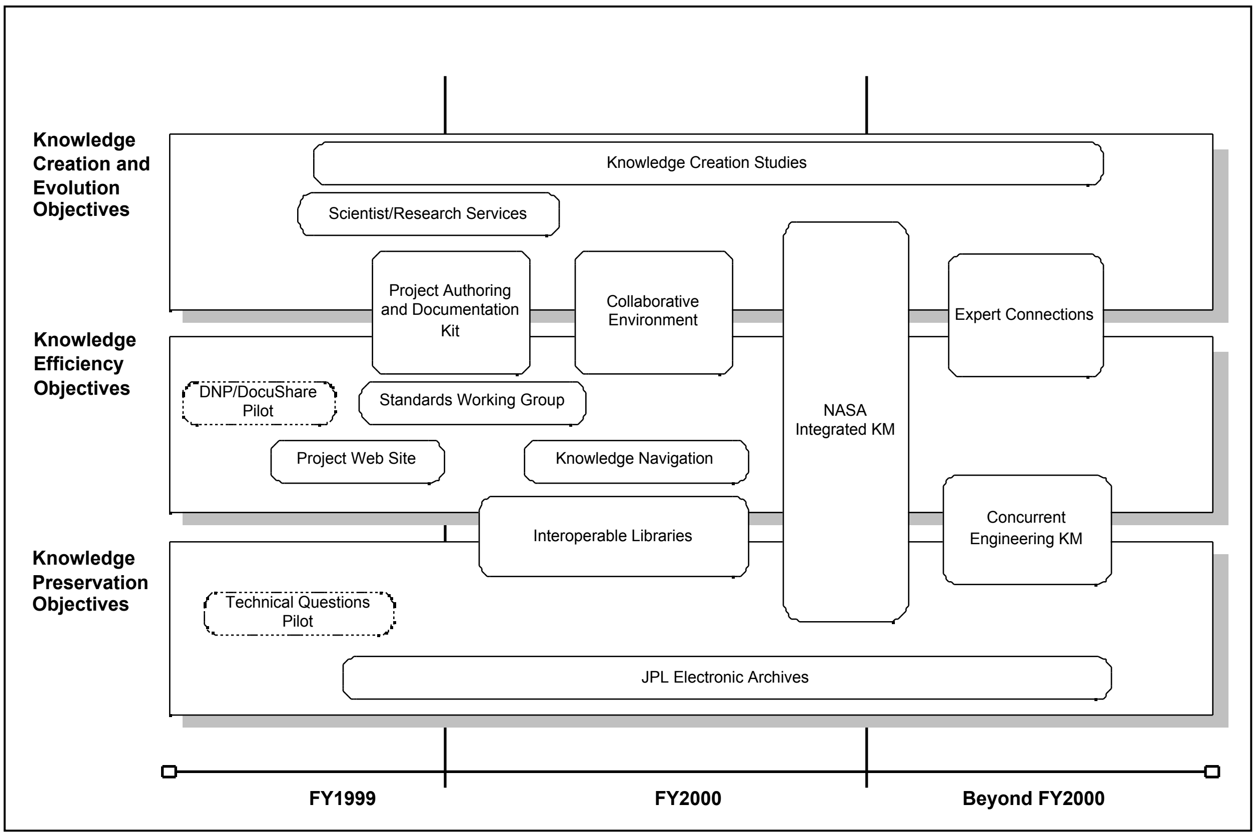


Figure 6-1. Phased Implementation of Knowledge Management Initiatives

6.3.1 Project Web Site

Goal

- Reduce development and maintenance costs and improve effectiveness of knowledge sharing internally within projects by creating standards for project Web sites

Tasks

- Create templates for standard project Web interfaces and site structures and instructions for use
- Create standard Web configurations and publication processes to ensure security and policy compliance
- Create processes and tools to separate project Web site *content* from *delivery mechanisms* to allow project staff to focus their contributions on their audience
- Recommend standard procedures to link Web sites to SQL databases
- Provide projects with help and knowledge in the form of
  - Web development experts available to answer questions
  - Consolidated services for Web site development and management
  - Ongoing presentations and training in Web site communications
  - Online tutorials, tools, and templates for building Web sites
- Participate in JPL and NASA Web groups to evolve consistent standards, tools, and service structures
- Create a Web site "closure" process, including identification of what should be archived, in what manner, and provision of processes and tools to get it done
- Provide measurements for project Web site use, growth, and customer satisfaction
- Provide description of skills needed to maintain service base

Partners

- Web developers from Sections 644, 393, and 389
- JPL Webmasters from other organizations
- DNP
- SESPD projects

6.3.2 Project Authoring and Documentation Kits

Goal

- Help projects reduce costs and schedule to complete required documentation and fulfill all regulations by reusing prior knowledge of documentation products.

Tasks

- Develop knowledge resource consisting of templates, guidelines, and processes for completing project documentation
- Create standard documentation lists (document trees) for projects in their various phases
- Develop templates, wizards, and forms for creating basic project documentation using JPL standard core products (such as Word, PowerPoint, and HTML editor) that include boiler-plate wording
- Provide examples from other approved documents
- Provide integrated support services for authoring, including editing, imaging (photo, video, and graphics), and printing
- Incorporate into procedures and templates the requirements levied by NPG 7120.5a, ISO, ITAR, and other policies and regulations pertinent to the Lab (Appendix F)
- Provide measurements for project documentation kit use, growth (kit components), and customer satisfaction
- Provide description of skills needed to maintain service base

Partners

- Develop New Products/Provide Leadership Process (DNP/PLP)
- Define and Maintain the Institutional Environment (DMIE)
- Proposal Center
- Technical Information Section (644)
- Project members

6.3.3 Interoperable Libraries

Goal

- Enable integrated search capability across multiple project and institutional libraries for transparent access to documents and data

Tasks

- Provide standards for security groups, and naming conventions across libraries
- Provide support services and standard processes for document and data publishing
- Provide standards for metadata assignment, document control, and access control
- Enable standard in-process archival assessments
- Evaluate design options and develop standard APIs or metadata exchange mechanisms to enable transparent data access across libraries
- Provide standard measurements for individual library use, growth, and customer satisfaction
- Build upon work by the DNP/DocuShare Pilot (Appendix E)
- Provide measurements for use of integrated search and retrieval methods
- Provide description of skills, tools, and facilities needed to maintain service base

Partners

- DNP/PLP/Project Information Management (PIM)
- DMIE
- PDMS
- EIS Data Access
- JPL Technical Library

6.3.4 Collaborative Environment

Goal

- Improve meeting collaboration with foreign and domestic partners on projects so that all project members can more easily get the information they need, communicate with other project members, and work remotely

Tasks

- Implement NASA core software standards in prototype desktop and conference room environments
- Provide fully configured laptop computers for travel checkout
- Deploy and support voice- and dataconferencing within JPL and with partners
- Provide meeting scribes and tools to capture important decisions as they are made in meetings
- Incorporate security regulations (including data access) within procedures and building blocks
- Standardize the use of virus scanning software
- Standardize access to JPL electronic information for distributed users
- Provide measurements for collaborative environment use and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- Enterprise Network and Telecommunications
- Contracts/Contract Technical Monitors
- JPL Facilities
- DNS Alliance
- ICIS Customer Services
- Collaborative Engineering Environment Group
- Audiovisual Services

6.3.5 Expert Connections

Goal

- Let people easily find the experts or expertise (in-house or outside) that they need

Tasks

- Develop an experts' directory that provides basic information about how to contact experts in various areas and strives to capture some of the expert's knowledge (see Appendix E, Technical Questions Knowledge Base)
- Support the Element Scientists, Principals, and Technology Community Leaders in creating communications (interest groups, forums, brown bags, presentations, meetings, newsgroups, and electronic threaded discussions)
- Work with Human Resources to streamline hiring of new JPL employees and consultants (including creating job descriptions for "experts" in various disciplines)
- Provide measurements for directory and communications use, growth, and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- Human Resources
- Caltech
- Technology Affiliates Office
- Program and Project scientists
- Chief technologists
- JPL Technical Library

6.3.6 Standards Working Group

Goal

- Assure that JPL meets or exceeds NASA, industry, and academia standards for knowledge management processes and technologies

Tasks

- Identify standards working groups that are integral to knowledge management (see Appendix B)
- Appoint people working in the knowledge management services to serve on standards committees within industry and on NASA intercenter working groups
- Create in-house standards groups to resolve ongoing issues related to metadata standards, desktop recommendations and support, data categorization and taxonomies, data exchange formats and protocols, and standard APIs
- Develop core metadata standards for use across repositories and applications
- Begin work on a Knowledge Base data dictionary to help people identify terms across multiple repositories
- Provide measurements for standards use, growth, and customer satisfaction

Partners

- NASA CIO Standards Working Groups
- JPL Technical Library
- DNS Alliance
- Design Hub
- Standards Group
- Chief Engineer

6.3.7 JPL Electronic Archives

Goal

- Provide a JPL-wide electronic archive that allows people to get easy access to institutional information and provides projects a centralized one-stop shop for transfer of inactive and project documents, data, and records requiring long-term archiving

Tasks

- Identify existing institutional resources of information (such as Engineering Documents and Drawings/Vellum Files, Archives, and Records Center) (see Appendix H)
- Implement updated archive policies from ISO working groups
- Develop procedures for in-process assessment of information developed by the project throughout its life
- Describe procedures (including service providers and funding sources) for projects to archive their knowledge at the end of the project
- Develop and maintain an archival document and data repository for electronic and/or hardcopy storage for documents and drawings under various control levels
- Provide measurements for archives use, growth, and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- ISO Working Groups (Document and Data Management and Records)
- JPL Archives
- Data Distribution Lab
- TMOD
- PDMS

6.3.8 Knowledge Navigation

Goal

- Create an enterprise Web gateway to JPL's knowledge resources and easily customizable personal and group Web sites for easy access to institutional information and targeted delivery of information requested by individuals or workgroups

Tasks

- Perform internal evaluation of existing Electronic Lab-wide Information Access Site (ELIAS) Web site
- Design and implement changes to ELIAS to become the enterprise entry point (or *portal*) into JPL's knowledge and information resources
- Provide easy tools and services for individuals and groups to customize their own JPL Web space
- Provide easy searching of and access to JPL-owned (or subscribed) knowledge and information resources
- Allow people to search across multiple repositories (at the metadata level at least)
- Provide single user authorization for ease of accessing restricted information
- Support individual and group subscriptions to internal and external electronic information
- Participate in NASA Webmasters working group
- Provide measurements for use, growth, and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- NBS/IBS
- Electronic Communications Committee
- Division 64
- EIS
- ELIAS
- ICIS

6.3.9 Knowledge Creation Studies

Goal

- Improve quality, methods, and rate of capturing the knowledge that JPL employees create

Tasks

- Continue to ask JPL employees what motivates them to share knowledge and what knowledge they find useful (feed that information back to the other tasks)
- Create incentives for contribution, use, and reuse of knowledge at JPL
- Develop classes, references, a help desk, and communications in how to use and contribute to JPL's knowledge
- Benchmark with industry and academia to find innovative ways of helping people move knowledge into a tangible format that others can use
- Develop a portfolio of knowledge management techniques (both organizational and technical) and their known or expected relationships with knowledge management objectives
- Provide measurements for knowledge creation, reuse, and customer satisfaction with knowledge resources
- Provide description of skills needed to maintain and extend service base

Partners

- Academic knowledge management consultants
- Division 31 measurement experts
- Primary process domain owners
- Human Resources
- Generate Scientific Knowledge members

6.3.10 NASA Integrated Knowledge Management Environment

Goal

- Determine costs and benefits of deploying robust integrated COTS tool suite as knowledge management framework (Livelink)

Tasks

- Evaluate other NASA Centers' use and customization of Livelink
- Implement integrated Livelink features (e.g. calendaring, forms, or workflow to support process improvement
- Incorporate security regulations (including data access) within procedures and building blocks
- Standardize the use of virus scanning software
- Standardize access to JPL electronic information for distributed users
- Provide measurements for integrated knowledge management system use, growth, and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- Collaborative Engineering Environment Group
- Proposal Center (HATDAM)
- NASA centers
- Enterprise Network and Telecommunications

6.3.11 Concurrent Engineering Knowledge Management

Goal

- Improve concurrent engineering processes by providing standard design structures, interchange processes and tools, and interfaces and increasing reuse potential with automated archiving

Tasks

- Review NASA standards for engineering design interoperability and prioritize projects and/or discipline teams to begin use
- Identify most frequent data conversions needed, and develop and/or procure tools
- Design generic templates and standard interface definitions for designs and models
- Design and integrate data exchange processes into DNP processes
- Integrate engineering data repositories into institutional Product Data Management (PDM) processes and systems

Partners

- Design Hub
- NASA CAD/CAE standards Working Group
- Industry Partners (e.g. Lockheed-Martin)
- Engineering Data Management Group (EDMG)
- Chief Engineer
- Applied Technology Program

6.3.12 Scientist/Research Services

Goal

- Design and integrate expert research services into the early phases of new projects and proposals to extend core JPL knowledge centers (e.g. Centers of Excellence)

Tasks

- Evaluate current methods of researching topics and ideas for new JPL investigations and partnership opportunities
- Determine specific research needs of JPL Centers of Excellence
- Identify internal and external resources and institutional capabilities to improve research methods
- Develop expert research methods and processes, and integrate with proposal process
- Provide measurements for expert research service use, proposal process improvement, and customer satisfaction
- Provide description of skills needed to maintain and extend service base

Partners

- JPL Library
- Earth and Space Science Division
- Centers of Excellence (In Situ pilot)
- Proposal Center
- Team X or A

6.4 Success Criteria

It became clear in looking across the literature and in performing the benchmarking that there were critical success factors across organizations in implementing knowledge management (see Figure 6-1). These success factors can be applied to JPL in four specific areas to determine how well JPL's knowledge management efforts adhere to known success criteria during implementation. The four areas involve JPL culture, the knowledge architecture, services and tools infrastructure, and the information technology infrastructure.

6.4.1 JPL Culture

In the area of cultural issues, critical areas include ownership, sharing and reuse of knowledge, and the incentives and rewards JPL gives to people who contribute and reuse knowledge. Specifically, the implementation effort should help JPL move to an environment where

- Sharing and reuse of JPL- and externally generated knowledge is a JPL core competency that enables faster-better-cheaper performance in each of JPL's endeavors
- JPL employees are rewarded and recognized for sharing and reusing knowledge
- JPL treats knowledge as a reusable commodity

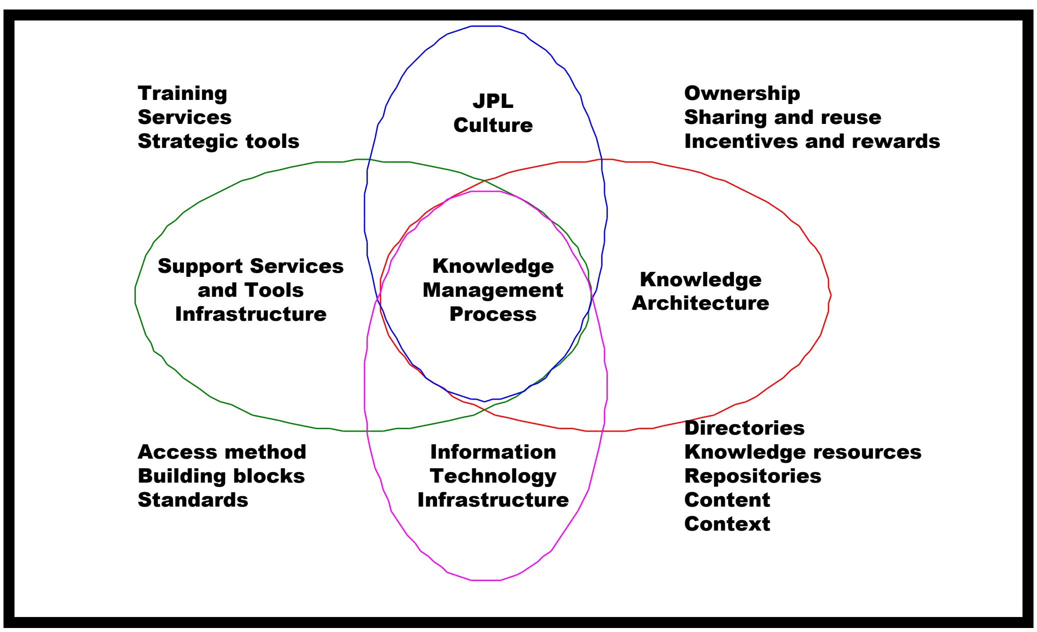


Figure 6-1. Critical Success Factors in Knowledge Management

6.4.2 Knowledge Architecture

In the area of knowledge architecture, critical areas include directories, knowledge resources, repositories, and the content and context of the knowledge. Specific architectural issues should strive to

- Internal and external sources of information and listings of experts are easy to find
- Content and context are the responsibility of the people who develop the knowledge (the knowledge providers). In creating knowledge, providers flag important issues for secondary users: issues such as the accuracy, inheritance, and use of the information
- Knowledge that is critical to JPL's business is the responsibility of "stewards" who ensure that the information is kept up-to-date and easy to use

6.4.3 Supporting Services and Tools Infrastructure

In the area of supporting services and tools infrastructure, training, services, and strategic tools are key to success. Implementation should focus on ensuring that

- JPLers can get information at their or team workstations, conference rooms, at home, or on travel
- JPLers have access to information they need online (including help desk level support), by request from a research service, or through normal communication
- No user manuals are required to navigate JPL's knowledge, but training and support is available when requested
- Services exist to help JPL employees as they create, develop, organize, and distribute knowledge
- If knowledge providers create information in a standard format, then tools are available for others to convert that information into any other standard format

6.4.4 Information Technology Infrastructure

In the final area of information technology infrastructure, success has been shown to depend on the access methods, building blocks, and standards that are implemented, such as ensuring that

- The system architecture provides the capability to restrict the availability of certain JPL knowledge to specific groups of people
- An open system architecture imposes minimal standards to ensure usability of the knowledge management capabilities internal and external to JPL, while permitting knowledge providers maximum flexibility. This architecture should be open-ended in terms of capacity and knowledge types and be able to gracefully accommodate changes.

6.5 Implementation Plan Considerations

6.5.1 Existing Process and Services Base

A great deal of knowledge management is already performed at JPL. Processes and services currently support knowledge management related functions. The forthcoming *A Knowledge Management Implementation Plan for JPL* will identify, to a first level, the existing processes and services that parallel recommended knowledge management services. The intent is to expand and include existing services, rather than to reinvent service bases. (The intent for processes is to incorporate existing processes into the knowledge management process structure proposed in Section 3. Those affected processes will be identified in the *Implementation Plan*.)

6.5.2 Organization

The Study Team recommends that funding for knowledge management be entrusted to ICIS and distributed to in-house sections, consultants, and JPL partners to accomplish specific objectives in relation to the pilots and underlying infrastructure needed to make knowledge management succeed.

A Knowledge Management Task should be created in Office 174, Enterprise Information System Applications. The Knowledge Management Task Manager should be located within Office 174, while the Work Package Managers, knowledge management process owners, and support staff are distributed across the line organizations that are or will be supporting knowledge management.

6.5.3 Partnering Goals

The Implementation Team will need to partner with other internal and external organizations to achieve success. In addition to partnering with current internal knowledge management related process owners and service providers, the Implementation Team needs to partner with

- JPL projects and programs in various lifecycles
- JPL institutional services
- JPL process owners
- NASA community
- NASA Chief Information Officer
- Industry leaders in knowledge management for additional lessons learned
- Industry partners

6.5.4 Budget Consolidation Opportunities

As different knowledge resources are identified, it may become apparent that duplicative work is being performed under two separate tasks or by two different organizations. In cases where two groups are performing tasks or providing services that can be combined, the Implementation Team should seek to consolidate those operations, particularly if both groups are funded out of JPL burden.

7 Knowledge Resources

The knowledge management architecture is designed to provide a standard mechanism for creating, improving, locating, and using a variety of JPL knowledge bases. The following list offers some examples of knowledge resources in several areas. It is not meant to be a comprehensive list, but instead to provide a context for the kinds of information that may be put into a Knowledge Base.

7.1 Business and Employee Resources

JPL is in the process of a major conversion to a commercial suite of core applications as part of the New Business Solutions (NBS) re-engineering effort. The databases underneath these new applications should be considered the "gold source" of data for information for core human resources, financial, and supply-chain business processes. The business processes that surround this suite of applications must ensure the accuracy and retention of this content.

7.2 Technology Resources

Developing a new technology takes effort; however, exceptional effort is often required to install that same technology into a JPL product or process. Aside from high costs, convincing someone to risk something new rather than being safe with an old (less capable) standby is very difficult.

Formal representation of the capabilities, costs, and resource properties of candidate technologies can substantially reduce resistance to technology transfer, taking advantage of: feedback on strengths and weaknesses; debriefings on outcomes; formalized communication between projects and researchers or commercial partners; or even a technologist database to communicate people's expertise.

7.3 Research Services

Research services to help meet research needs are offered by the JPL Technical Library and other facilities, including outside providers. Much of the current research capability resides in the technical staff and experts within their own fields. Formalized research services primarily focus on information search and retrieval and collection management. The speed and effectiveness of these services could be enhanced in several ways:

- Full-text search and retrieval for selected documents
- Improved access to legacy JPL-produced documents
- Standardized metadata and online search capability

- Improved consistency of cataloging and metadata representation across JPL collections of current documents
- Common representations for internally and externally produced documents
- Utilizing research service providers at JPL, such as the research librarians in Section 643

7.4 Science Resources

The primary objective of the missions conducted by JPL is to collect and disseminate science data and investigation results to the scientific and educational communities and to the general public. Data received from both Earth and space science missions is processed at several levels and made available to the appropriate community for further analysis. The science community has strong peer review mechanisms in place, and in many ways is a model for the manner in which effective knowledge management systems can and should operate. At JPL, a number of science data systems support these knowledge dissemination efforts:

- *Planetary Data System*—composed of a distributed set of academic institutions and service groups nation-wide, each with expertise in a specific area of planetary science
- *Physical Oceanography DAAC (PO.DAAC)*—supplies science data from JPL's Earth science missions
- *Alaska SAR*—provides data from a variety of radar experiments
- *Multimission Image Processing Lab (MIPL)*—provides imaging data to industry and the public

While much of this data is produced at JPL, there is no institutionally managed catalog of these data sets. As a result, determining consistent and reliable metrics for the quantity and quality of data actually produced and disseminated from JPL-managed missions is a difficult task.

7.5 Product Development Resources

Product development activities have substantial knowledge management requirements. Management processes and tools needed for cost-effective and successful product development and operations require correlation of several different data sources. Decisions requiring trades between costs, objectives, designs, methods, operations, and risk require human resources information, estimates, spacecraft design documents, schedules, statistical analyses, parts lists, supplier performance assessments, operations scenarios, historical data, and floor plans.

7.6 Infrastructure Service Resources

There are many infrastructure services at JPL, which provide the skilled personnel and physical or electronic resources necessary to support its primary missions and projects. Some examples of these services and the types of knowledge resources they contribute to the JPL institutional knowledge resource base include:

- Education and training—catalogs containing courses and descriptions, training records, and online training materials
- Reproduction—catalog of reproduction satellites, hours, and fees; electronic submission forms
- Enterprise Information System—service descriptions and fees, procedures for signing up and reporting problems
- Technical Library—targeted research help, book ordering procedures, and online research materials
- Electronic publishing—documentation services and templates; and writers and editors available
- Photo Lab—information on still and video shoot capabilities; electronic and print image services; and electronic ordering and archives
- Graphics—graphics catalog and repository; and print capabilities
- Audiovisual—catalog of videotapes and services, narrators and sample voices, and multimedia production rates and schedules

Many of these services already provide and maintain databases accessed and used internally or by the JPL community. The knowledge management architecture would enable these resources to be improved and allow discovery, search, and access to them in a standard way.

7.7 JPL Customer Resources

JPL has a variety of customers—including NASA, Congress, other government agencies, scientists, industry, the educational community, industry, and the public. Many offices within JPL are dedicated to customer outreach, including Educational Affairs, Public Affairs, International Affairs, and Commercial Technology Program Office.

Educational and public outreach efforts have become an increasingly important part of JPL’s focus, resulting in many exciting educational materials generated as a result of individual projects’ creative responses to the demand for greater communication, accountability, and visibility. Unfortunately, there is no institutional resource base or maintenance service for products developed outside the auspices of the primary customer service offices. As a result, it is again difficult to track exactly what JPL has disseminated in the way of educational and public information. It is also difficult to determine the effectiveness of these materials.

Knowledge resources geared to each audience should be available in a standard way and presented according to themes, targets or subject matter, in addition to providing the material in the context of a specific project. NASA's move to four major science themes recently has provided a good framework for this.

8 References

[1] *AIIM International '98 Conference Handbook.* Association for Information and Image Management International, May 11–13, 1998: http://www.aiim.org

[2] Alavi. M. and D. Leidner. "Knowledge Management Systems: Emerging Views and Practices from the Field," 1998: http://thorplus.lib.purdue.edu/library_info/departmental_papers/1582.html

[3] Allee, V. "Chevron Maps Key Processes and Transfers Best Practices", April 1997: http://Webcom.com/quantera/Chevron.html.

[4] Amoedo, J., R. C. Eddington, J. R. Kahr, P. M. B. Shames, R. A. Somer, and J. S. Jenkins. *Enterprise Information System Architecture*. JPL D-12991 (internal document), Jet Propulsion Laboratory, California Institute of Technology, November 1995.

[5] "Argument-Based Design Rationale Capture Methods for Requirements Tracing"; "Component-Based Software Development/COTS Integration"; "COTS and Open Systems—An Overview"; "Distributed/Collaborative Enterprise Architectures"; "Domain Engineering and Domain Analysis"; "Middleware"; "Object Request Broker"; "Organization Domain Modeling"; "Requirements Tracing—An Overview"; *Software Technology Review*: http://www.sei.cmu.edu/str/

[6] Carpi, C. *Review of DMIE Document Management*, Information Architects, Version 1, (internal document), Jet Propulsion Laboratory, California Institute of Technology, September 23, 1998.

[7] *Communicating NASA's Knowledge*, NASA NP-1998-08-240-HQ, National Aeronautics and Space Administration, Washington, DC, August 1998.

[8] *Computer-Aided Engineering, Design, and Manufacturing Data Interchange Standards*, *Draft*, NASA-STD-2817, National Aeronautics and Space Administration, Washington, DC.

[9] *Content Management: Putting Knowledge to Work*, Washington, D.C. Special Libraries Association Conference, November 5–6, 1998.

[10] Cranford, S. "Knowledge Through Data Warehousing," *DM Review*, September 1998.

[11] Davenport, T. "Hewlett-Packard Promotes Knowledge Management Initiatives", June 1996: http://Webcom.com/quantera/HP.html

[12] Davenport, T. and L. Prusak. *Working Knowledge: How Organizations Manage What They Know.* Harvard Business School Press, 1998.

[13] Davenport, T. H. "Some Principles of Knowledge Management": http://www.bus.utexas.edu/kman/kmprin.html

[14] Davenport, T.H., "Knowledge Management Case Study: Knowledge Management at Microsoft, 1997": http://kman.bus.utexas.edu/kman/microsoft.htm

[15] Delio, M. "Keys to Collaboration," *Knowledge Management*, October 1998: http://www.kmmag.com

[16] Drucker, P. F. "The Future That Has Already Happened," *Harvard Business Review*, September–October 1997.

[17] Frew, J., M. Freeston, et al. "The Alexandria Digital Library Architecture", Lecture Notes in Computer Science 1513, *Research and Advanced Technology for Digital Libraries Proceedings*, Second European Conference, ECDL '98 Heraklion, Crete, Greece, September 1998.

[18] *Guidance for NASA Public Key Infrastructure (PKI), Draft Version 1.2.4*, National Aeronautics and Space Administration, Washington, DC, November 16, 1998.

[19] Hars, A. and A. Malhotra. *"Emerging Web-Based Knowledge-Bases: Categories, Trends and Implications for the IS Research Community."* University of Southern California, Los Angeles, 1998: http://www-rcf.usc.edu/~hars/cybrarium/survey1.html

[20] "IBM DB2 Digital Library": http://www.software.ibm.com/dug-lib/about.htm

[21] *IBM Open Blueprint, Digital Library Resource Manager*, GC23-3909-01, IBM Corporation, 1996.

[22] *Information Technology Implementation Plan FY1999-2003*, National Aeronautics and Space Administration, Washington, DC, February 1998.

[23] Inmon, W. H., C. Imhoff, and R. Sousa. *Corporate Information Factory.* John Wiley & Sons, Inc., 1998.

[24] Inmon, W. H., J. A. Zachman, and J. G. Geiger. *Data Stores, Data Warehousing, and the Zachman Framework: Managing Enterprise Knowledge.* McGraw-Hill Publications, 1997.

[25] *Interoperability Profile for NASA E-mail Clients*, NASA-STD-2808A, National Aeronautics and Space Administration, Washington, DC, January 23, 1998.

[26] *Intracenter Networking Architecture, Standards and Products*, NASA-STD-2802, National Aeronautics and Space Administration, Washington, DC, May 1, 1997.

[27] *JPL Automated Information Security Requirements for Computer Systems Administrators*, (internal document) Jet Propulsion Laboratory, California Institute of Technology, November 1996.

[28] Kimball, R., L. Reeves, M. Ross, and W. Thornthwaite, *The Data Warehouse Lifecycle Toolkit*. New York, NY: John Wiley and Sons, Inc., 1998.

[29] "Knowledge Discovery from Data and Data Mining", MITRE: http://www.mitre.org/pubs/showcase/datamining/index.html

[30] *Knowledge Leadership Study*. The Delphi Group. Boston, MA, 1998.

[31] Majchrzak, A., N. King, R. E. Rice, A. Malhotra, and S. Ba "Computer-Mediated Knowledge-Sharing and Re-use Among Members of a Creative Inter-organizational Virtual Team", submitted to *Management Information Systems Quarterly*, 1998.

[32] Manasco, B. "Dow Chemical Capitalizes on Intellectual Assets", March, 1997: http://Webcom.com/quantera/Dow.html

[33] Manasco, B. "Steelcase Designs the Intelligent Workspace", August 1996: http://Webcom.com/quantera/Steelcase.html

[34] Manasco, B. "Sun's Knowledge Network Enhances its Selling Skills", May 1997: http://Webcom.com/quantera/Sun.html

[35] March, A. "A Note on Knowledge Management", Harvard Business School 9-398-031, November 26, 1997.

[36] McGee, J. and L. Prusak. *Managing Information Strategically.* John Wiley & Sons, Inc., 1993.

[37] McHugh, J. "Understanding Repositories, Putting Metadata to Work", *Distributed Computing*, July 1998: http://www.DistributedComputing.com

[38] *Minimum Hardware Configurations, Draft*, NASA-STD-2805, Rev. 2, National Aeronautics and Space Administration, Washington, DC, March 31, 1998.

[39] *Minimum Office Automation Software Suite Interface Standards and Product Standards, Draft*, NASA-STD-2804, Rev. 2, National Aeronautics and Space Administration, Washington, DC, March 31, 1998.

[40] Morey, D. and T. Frangioso. "Knowledge Management Systems," MITRE presentation, MITRE, September 24, 1997: http://www.mitre.org/resources/centers/it/g062/kmsystems/1997review/index.html

[41] Myers, P. S. and R. W. Swanborg. "Packaging Knowledge," *CIO Enterprise*, April 1998.

[42] *NASA Directory Service, Architecture, Standards and Products*, NASA-STD-2816, National Aeronautics and Space Administration, Washington, DC, in process.

[43] *NASA Electronic Messaging, Architecture, Standards and Products*, NASA-STD-2815, National Aeronautics and Space Administration, Washington, DC, in process.

[44] *NASA Firewall Strategy, Architecture, Standards and Products*, NASA-STD-2813, National Aeronautics and Space Administration, Washington, DC, August 19, 1997.

[45] *NASA Integrated Information Technology Architecture*, NASA-STD-2814, National Aeronautics and Space Administration, Washington, DC, September 9, 1997.

[46] *NASA Intelligent Electronic Forms*, NASA-STD-2809, National Aeronautics and Space Administration, Washington, DC, April 10, 1997.

[47] *NASA Strategy for Windows NT Domain*, NASA-STD-2801, National Aeronautics and Space Administration, Washington, DC, May 1, 1997.

[48] *NASA Technical Standards Program Infrastructure and Products*, National Aeronautics and Space Administration, Washington, DC, October 5, 1998.

[49] *NASA X.500 Directory*, NASA-STD-2807A, National Aeronautics and Space Administration, Washington, DC, February 20, 1996.

[50] Nonaka, I. And H. Takeuchi. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, 1995.

[51] Payette, S. and C. Lagoze. "Flexible and Extensible Digital Object and Repository Architecture (FEDORA)", Lecture Notes in Computer Science 1513, *Research and Advanced Technology for Digital Libraries Proceedings*, Second European Conference, ECDL '98 Heraklion, Crete, Greece, September 1998.

[52] Pinelli, T. E., R. O. Barclay, J. M. Kennedy, and A. P. Bishop. *Knowledge Diffusion in the U.S. Aerospace Industry*, Greenwich, Conn.: Ablex Publishing Corporation, 1997.

[53] *Principle Center for Workgroup Computing (HW and SW),* Task Summaries and Status, Lewis Research Center, Cleveland, OH, April 8, 1998.

[54] Raitt, D., S. Loekken, J. Scholz, H. Steiner, and P. Secchi. "Corporate Knowledge Management and Related Initiatives at ESA", November 1997: http://esapub.esrin.esa.it/bulletin/bullet92/92raitt.html

[55] Robie, J. "XML and Modern Software Architectures", Texcel Ventures, Inc., 1997: http://www.texcel.no/

[56] Santosus, M. “Information Micromanagement,” *CIO Enterprise*, April 1998.

[57] *Seven Steps to Implementing Knowledge Management in Your Organization*. Dataware Technologies, Inc., 1998: http://www.dataware.com

[58] Sprague, R. H. “Electronic Document Management: Challenges and Opportunities for Information Systems Managers,” *MIS Quarterly*, March 1995.

[59] Targowski, A. S. *Global Information Infrastructure.* Harrisburg, PA: Idea Group Publishing, 1996.

[60] *The Enterprise Architecture Conference Proceedings*,Toronto, Ontario, DCI, Andover, MA, May 26–28, 1998.

[61] *Treasury Information System Architecture Framework, Version 1.0*, Office of the Deputy Assistant Secretary for Information Systems and Chief Information Officer, January 3, 1997: http://www.ustreas.gov/tisaf/

[62] Ulrich, D. “Intellectual Capital = Competence × Commitment,” *Sloan Management Review*, Winter 1998.

[63] Vlahos, W. *Virtual Private Networks (VPN) Functional Requirements Document (FRD*) Version 1.0, Draft Revision 5, Jet Propulsion Laboratory, California Institute of Technology (internal document), December 15, 1998.

[64] Walsh, N. “What is XML”: http://www.xml.com/xml/pub/98/10/guide1.html

[65] Wang, R. Y., Y. W. Lee, L. L. Pipino, and D. M. Strong. “Manage Your Information as a Product,” *Sloan Management Review*, Summer 1998.

[66] *Workgroup Hardware/Software Architecture, Standards and Products Preview—Draft*, National Aeronautics and Space Administration, Washington, DC, December 4, 1998:

A Benchmarks

A.1 Who's Doing Knowledge Management?

The expression often cited in the literature of why knowledge management was first investigated is "*If we only knew what we know.*" Many companies have implemented knowledge management. The KM Study Team chose to benchmark with the ones below, which are sorted as to the type of benchmarking method through which the information was obtained.

- Site Visit
  - The AeroSpace Company
- Telephone Interview
  - Hughes Space and Communications
- Case Studies
  - AutoSTEP
  - Ernst and Young
  - Hewlett-Packard
  - Microsoft
  - MITRE

The benchmarking included a site visit, interviews[1], analysis of published case studies, and consulting with academic leaders in knowledge management. In addition, to extensive literature research and sponsoring attendees at conferences, the Team's goal was to focus not on just the technology aspects, but on understanding key human and organization factors. The outcome on each benchmark was to understand the scope of their implementation, evaluate the maturity of the tools, define the processes, and describe the quantifiable results obtained.

A.2 Case Studies and Benchmarks

A.2.1 The Aerospace Corporation

The Aerospace Corporation is a private, nonprofit corporation that conducts research and development, and advises aerospace industry companies. Most of the company's work is related to the design, test, evaluation, and initial operation of space systems. To provide a full service for the aerospace industry, Aerospace uses a matrix of 1000 specialists in science and engineering; each expert supports nearly 50 programs concurrently. Technical learning and knowledge sharing occur among the experts. Aerospace built a technical infrastructure that provides a carefully designed set of tools, including engineering models, simulations, and laboratory facilities.

[1] The survey instrument is included as Figure A-2 at the end of this Appendix.

The Aerospace Corporation started with knowledge management from the conceptual design phase. The primary goal of the project is to provide a common interface to information resources and to interconnect separate databases. They had a rapid prototyping, three-phase implementation schedule: phase one was for collection and indexing; phase two was interactive user participation, such as publishing and pushing; and phase three is currently information refinement.

Aerospace focused on their main processes, and built an information infrastructure to share technical knowledge. They have many critical and important defense contracts, so they employ a layered security approach using an enterprise firewall, virtual private networks, access control lists, and password authentication. The cost effectiveness and sharing of knowledge are very important to the company. These factors encourage their reuse and sharing culture, and they attribute knowledge management's success to this cultural factor.

A.2.2 Hughes Space and Communications Company[2]

Hughes Space and Communications Company (HSC) is a unit of Hughes Electronics Corporation and is the largest commercial satellite manufacturer in the world. Since 1961, HSC has engaged in the development and production of space and communication systems for military, commercial, and scientific uses. Because of their industrial position, they have high risks and a sizable investment. Due to these factors, HSC tried to cut down the costs of building satellites as much as possible. External market competition has driven constant technology changes, along with the need to provide a broad-spectrum service and to handle customer requirements. HSC has a strategy to build its technology and experience to develop new applications for its satellite services. HSC seeks the opportunity to maintain their superior position in the satellite manufacturing market and to increase the profitability through more efficient production processes.

HSC does not view knowledge management as a simple process, a function, or an organization. Knowledge is a skill to manage their business and an essential tool for a manager. HSC does not direct the project, but provides services to people to enable them to do their work better. The system was started from the top, rather than in the individual business units. The overall goals of the system are to increase efficiency and cut down the satellite-building cost. To provide valuable information to their top management, they integrated many databases within the corporation.

The knowledge management system extracts, records, and provides people's successes and failures. It connects what is learned with what is practiced. Lessons are documented and disseminated to employees. The feedback to those lessons is also collected and documented.

[2] The following six case studies were analyzed by Jongbok Byun, Claremont Graduate University.

HSC is implementing an intranet through a phased pilot approach. The pilots focus on building a robust system that supports a high business value resource, such as lessons learned, yellow pages, or a common user interface for an existing system. Using stepwise tutorials, each employee is trained in how to use the intranet for their job.

HSC used a spiral development methodology in knowledge management and built prototypes. The system is based on existing components that are familiar to users. The knowledge management team utilized existing, well-operating database functionality; still, many systems and components are running in the business units since HSC has a distributed information technology infrastructure. HSC is trying to build an overall corporate system based on internet standards, especially using Netscape. By placing information on the Web, people can easily access and modify their information. Dynamic reports are generated when people update the underlying data. Still much unrelated, unused, and unreliable data is located in the integrated system, and HSC plans on integrating data cleansing and more metadata standards.

The cultural factor was also important here. At HSC, the knowledge architecture is critical, because they have a variety of data formats and information repositories. The architecture for content, context, and directory of resources is the promising critical success factor.

A.2.3 MITRE

MITRE performs system engineering and integration work for the Department of Defense C^3I, and systems research and development work for the Federal Aviation Administration. MITRE has over 5000 employees and its Corporate Information Services maintains a collection of over 175,000 items.

The first year-long knowledge management program was started in October 1994 [29, 40]. At the midpoint of the project, they found that the role of the knowledge manager was very similar to the librarian. So this program offers a knowledge management model for a digital library environment. In the networked information environment, collection development, resource organization, bibliographic instruction, and reference services are the basic roles of the digital library.

MITRE supports multiple platforms, providing access to internal and external resources. In order to share knowledge among diverse and distributed groups, MITRE has combined new technologies with structured processes for publishing and organizing information. Currently, MITRE is engaged in the MITRE Information Infrastructure (MII) for the entire corporation. MII will serve as the information repository for all MITRE-generated documents.

MITRE recognized that the reengineering of the information environment has many effects on the company. In order to ease the transition and to identify the information needs of the technical staff, the knowledge manager position was created. The knowledge manager has the responsibilities of organizing information, creating metadata, navigating the data warehouse, creating information and people locators, providing context-rich information, using internet tools to manage information space, and defining locator concepts and operations to allow software knowledge agents to be built. The knowledge manager's team has built the internal and external information network by using internet tools. The knowledge manager participated in the design efforts of the corporate information infrastructure.

User education is one of the most important roles of the knowledge manager at MITRE. New employees are given MITRE online resources, as well as the external collections. In MITRE, great emphasis is placed on training and collection development, while resource organization is facilitated by the standard process establishment. The role of knowledge manager is dependent upon the organizational structure and information needs. MITRE notes that the *user* should be the major player in knowledge management, not the knowledge manager or the digital librarian.

A.2.4 Ernst & Young

As one of the largest financial and consulting companies, Ernst & Young (E&Y) is located in a knowledge intensive industry. They have many internal data resources and well-established consulting method repositories [35].

In 1993, the company launched a $1B strategic plan, "Future State '97" (FS '97), to get a competitive edge in the consulting business. This plan included the preferred future vision and plans of Ernst & Young's vision in processes, sales, service, delivery, people, and knowledge. The initial goals of knowledge management in Ernst & Young included capturing and leveraging knowledge from consulting engagements, having every consultant contribute to the firm's stock of knowledge, and becoming known by clients as an important source of knowledge and thought leadership.

Ernst & Young tried to speed up the processing time to provide consulting solutions to its clients. By 1995, the company used an approach called the Accelerated Solutions Environment, which included the rapid application of Ernst & Young knowledge, models, and approaches to client situations in facilitated large group settings.

E&Y appointed their first Chief Knowledge Officer for the FS '97 project. The CKO's main role was managing the overall processes and technologies of the firm that related to knowledge. He was supported by a Knowledge Process Committee in charge of the knowledge topics and means for their integrated knowledge management.

Organizationally, E&Y has created three important research centers since 1990

- The Center for Business Innovation was built for the creation of new knowledge and has researched issues such as business process reengineering, organizational change management, and knowledge management

- The Center for Business Technology was created to help structure knowledge into methods and automated tools to support the consulting business
- The Center for Business Knowledge (CBK) served as the library to gather and store both the firm's acquired knowledge and external knowledge and information. The CBK was transitioned to the central function of E&Y knowledge management. By the end of 1996, the CBK had more than 100 professionals, including a library, a call center for answering consultant requests, and a database of consultant skills

The CBK also had a Knowledge Management Network formed into key domains of knowledge within the consulting practice. Each Network had an online discussion and document database in Lotus Notes. Each Network was assigned a person half time to capture the knowledge from particular engagements, to prompt consultants to add their own learning, and to edit the discussion and document databases. The CBK was also responsible for the E&Y consultants' skills database. It established a new model for evaluating and describing consultant competencies. Each consultant was evaluated by his or her supervisor, and the information was entered into a database. Another important task of the CBK was developing a knowledge architecture and taxonomy.

In the beginning, E&Y opened their discussions and concerns about knowledge management. However, after the people in the organization became familiar with the knowledge management environment they wanted to set their standards. They structured and filtered their online materials including qualifications, sales presentations, proposal templates, and answers to frequently encountered issues. E&Y used Lotus Notes for their dispersed organization structures and knowledge base. Lotus Notes was selected as the primary technological platform for capturing and disseminating internal knowledge.

As with the knowledge architecture, E&Y allowed multiple technologies to proliferate in the early days of knowledge management. There were between 200 and 300 local applications and databases. Later, the firm wanted consultants to focus on content rather than applications. Approximately 12 to 15 applications eventually came to support knowledge management, including Notes, the Web, the skill database, and a few others. E&Y is now trying to standardize their hardware and software for knowledge management. These standards mean that programs and documents can be exchanged easily within the company. Using a common technology platform, E&Y is integrating all knowledge, model, tools, and techniques into their Accelerated Solutions Environment. Using ASE, E&Y's consultants get fast knowledge access. The CBK developed the knowledge objects, frameworks, and techniques used in it.

Still there are many organizational challenges to knowledge management. Embedding knowledge is still difficult. It was particularly difficult to use technology to support some types of consulting knowledge that is tacit, implicit, and difficult to extract from the minds of practitioners.

The cultural residence is also a barrier for the knowledge management. The traditional cultural environment of E&Y was pragmatic and situationally oriented. However, the knowledge management environment needs structured knowledge and applications. *One key means to enhancing this change was embedding a knowledge orientation into the firm's performance evaluation process, consultants are now evaluated in part on their contributions to and use of knowledge.* Another difficulty was the self-justification of knowledge management, since it is very difficult to measure the net economic value of the knowledge management process.

A.2.5 AutoSTEP

STEP is the new international ISO 10303 standard for neutral data formats. Using STEP, a company can exchange data between different brands of manufacturing software that deal with the same data. AutoSTEP is a product data exchange project for automotive supply chains. The goal of AutoSTEP is to improve the quality and timeliness of data exchange to minimize product-development cycle time. AutoSTEP focuses on both the STEP technology and product development business processes associated with product data exchange.[3] From the business view, AutoSTEP gives better integration of product and process design among automotive suppliers and customers. With AutoSTEP, the automotive industry gets accurate, timely, and cost-efficient exchange of product data and increases the coordination between development partners.

Table A-1. Applicability Matrix of STEP

	Design Responsibilities		
Design and Manufacturing Activities	Supplier	Joint	Customer
Design Packaging Analysis	Applicability of STEP (by combination of design responsibility with design or manufacturing activity)		

STEP has two key tools: (1) an applicability matrix (Table A-1), which is a guideline for when STEP is appropriate (the matrix is growing and destined to be a knowledge base for the automotive industry), and (2) a best practices guidebook that gives advice on improving product development practices. AutoSTEP defines nine practice issues for current problems such as CAD data quality issues and software version control issue. Based on the specific issue, a company can get possible solutions to improve their efficiency.

A.2.6 Microsoft

One of the greatest assets in Microsoft is its people. Microsoft prides itself on choosing the people most expert in specific technologies and businesses [14]. In November 1995, Microsoft started Skills Planning and Development (SPUD) for its IT employees'

[3] http://www.aiag.org/autostep/index.html

competency development, with the goal of transferring and increasing knowledge for IT personnel. SPUD consists of five main components

- Developing a structure of competency types and levels
- Defining competencies required for particular jobs
- Rating the performance of individual employees in particular jobs based on competencies
- Implementing knowledge competencies in an online system
- Linking the competency model to learning offerings

The entry-level competencies became a foundation of knowledge in the SPUD project.

As shown in Figure A-1, above the foundation level, there are local or unique competencies; these are advanced skills for a particular job. At the next level, global competencies mean those for a particular function or organization. The universal competencies are at the highest level and these are true for all employees in the company. Within each competency there are four levels of skill—basic, working, leadership, and expert.

With this view of four foundation competencies, there are two distinct domains. Explicit competencies are those for which there is knowledge and experience with specific tools or methods. Implicit competencies address abstract thinking and reasoning skills. In Microsoft, the implicit competencies are quite stable, but the explicit competencies change frequently due to the rapid environmental changes.

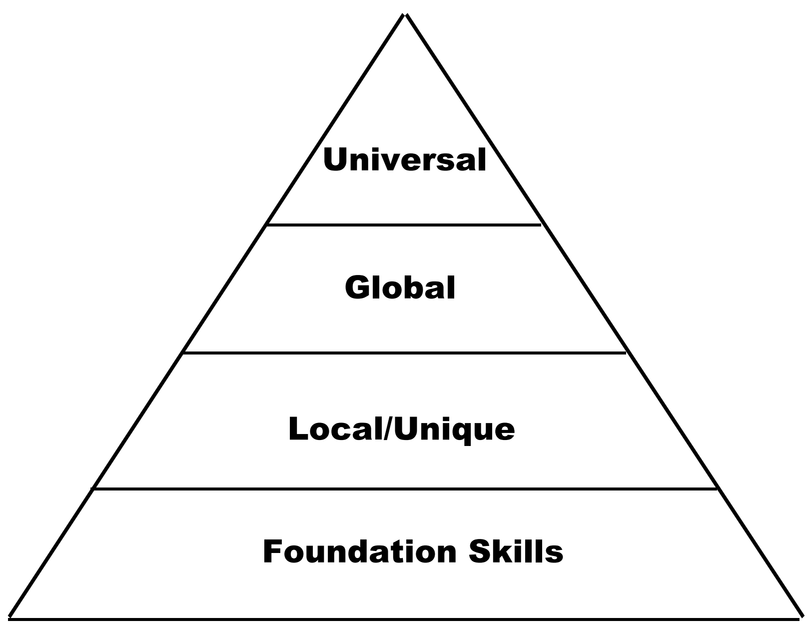


Figure A-1. Types of competencies at Microsoft

Using SPUD, Microsoft can find the best match between jobs and employee capabilities. Each Microsoft job has a required competency rating to which it must be performed. Evaluation of employees is another important role of SPUD. The overall goal of this rating is to build a competency inventory across Microsoft. A manager, using SPUD, can pick a qualified person without knowing his or her personal ability.

SPUD provides online competency structure, job rating, a rating database, and competency levels for employees. The SPUD has had technical success, but there are some problems to be solved. The job data needs to be managed centrally because it has personal information; however, there is no access control schema for the database, and data access permissions are a critical issue to Microsoft.

Microsoft wants to use SPUD for educational purposes. SPUD can link specific competency levels to specific educational courses and materials. Everything from internal education to external courses can be related to specific competencies and skill levels. The pilot test of SPUD has been well done, and they are trying to expand SPUD to a corporate level. One issue is how to integrate SPUD into the Microsoft product development framework because many employees are related to special product development.

Another issue is the relationship between SPUD and the overall human resources function. The human resources function provides some implicit competencies (e.g., team spirit) required throughout Microsoft.

The success of SPUD depends upon the behaviors of the individuals who use it. Employees and managers have to feel its usefulness. Finally, SPUD can advance knowledge by focusing on individual knowledge competencies that require active involvement by everyone in the organization.

A.2.7 Hewlett-Packard

The corporate culture of Hewlett-Packard (HP) is that of a distributed and self-regulated computer and electronics firm. There have been no top-down policies for knowledge management. However, in 1995, many divisions and departments began undertaking specific efforts to better manage knowledge by themselves. Many managers attempted to capture and distribute the knowledge resident in their own business units and departments. Some projects had several years' history and others were just beginning [11].

HP's corporate information systems group started a series of workshops to facilitate knowledge-management experience sharing within the company. The key objectives of the workshops were to facilitate knowledge sharing through informal networking and establish a common language and management framework for knowledge management. This is a "pull" approach to knowledge management, rather than a "push" strategy.

The first workshop was held in October of 1995. From this start, a list was compiled of than 20 HP sites where some form of proactive knowledge management was underway. The Corporate Education Group created a "Trainer's Trading Post," a discussion database on training topics that has also become a library of training materials and a means of evaluating training resources. A reward system for active participation in the post exists in the form of free Lotus Notes licenses and airline mileage.

The organizational culture of HP acts as *both* a facilitator and a barrier to knowledge sharing. HP has a relaxed, open culture that enhances knowledge exchange, and many technical engineers share their knowledge with each other. On the other hand, many department managers are reluctant to invest their time and money in knowledge sharing, because there is no obvious or immediate payback for their unit. The company's fast growth and change commonly lead employees to move from one business unit to another. This mobility helps informal knowledge transfer within the company. However, the company's decentralized organizational structure and operation mode work as a barrier to knowledge exchange. With a high degree of autonomy, there is little *organized* sharing of information, resources, and employees across units.

HP is developing a guide to human knowledge resources within HP Laboratories. The guide uses a Web browser and a relational database, and comprises expert profiles. These profiles point to the backgrounds and expertise of individuals who are knowledgeable on particular topics. The experts enter their own profiles and maintain them over time. However, it has proven to be very difficult to continue this without central assistance.

Many knowledge bases exist in Product Processes Organization (PPO), a corporate group with the mission of advancing product development and introduction. Competitor information, international marketing intelligence, and primary/secondary information are managed by the PPO. For example, one effort to capture and leverage HP product knowledge for the Computer Products Organization dealer channel, HP Network News, remarkably reduced the incoming call for technical support from dealers.

HP has internal expertise and external sources of knowledge within their knowledge base. Knowledge sharing over the knowledge management network has become a critical issue for successful knowledge management. HP wants to define who needs what kind of information. The workshops are one mechanism to understand who needs specific knowledge and how best to transfer it based on their pull policy for information.

It was unclear what steps are needed for a decentralized organization like HP to succeed in knowledge management. There are many small knowledge-sharing and knowledge-management practices, but there is no corporate level of support to facilitate knowledge sharing *between* business units. With the company's current success, it may be risky to change their processes at this point.

Currently, HP is trying to emphasize awareness building and developing a common vocabulary and frameworks through workgroups to help distinguish information management from knowledge management.

A.3 Analysis

The benchmarking has shown some clear trends in knowledge management implementations for phased implementation, innovative use of technologies, a need for security and standards, and large cultural issues unrelated to the method in which knowledge management is actually deployed.

A.3.1 Phased Implementation

Information systems generally are much too complex to consider monolithic and large-scale implementation, and this is especially true of knowledge management systems that are commonly aggregates of existing systems. Phased implementations using rapid prototyping is the approach that Aerospace Corporation is planning in its knowledge management system. Aerospace observed that the downside to rapid prototyping is declaring a prototype "operational" without including the robust features of a truly operational system. Due to the larger scope of knowledge management encompassing multiple information systems, a phased approach is clearly the only practical approach to implementation. Most initial implementations are focusing on collecting and indexing corporate knowledge.

A.3.2 Technologies

The knowledge management tool market is not mature and is very document-management oriented. *Modular standard-based components were shown to facilitate phased implementations.* One example of modular components is Java, which is fast becoming the language of choice along with Java-related technologies such as JNI (Java Native Interface) and Jini (a Java-based network object schema). Windchill, a new PDM toolset from Parametric Technologies, is a Java-based product that is being implemented at Lockheed-Martin and Airbus Industries.

A.3.3 Security

Information security issues are integral to knowledge management and must be woven into knowledge management processes and architectures. Users are typically not aware of all the security considerations related to the information they produce or access. Layered security is a common approach to scaling security in an appropriate manner and embedding security into the tools.

A.3.4 Standards

Cross-platform and data format issues can haunt the initial implementations of knowledge management. Standard workstation configurations are a best business practice. Effective use of standards is more about process than tools.

A.3.5 Culture

Incentives are crucial to the success of knowledge management and to changing the culture from one of knowledge hoarding to knowledge sharing. E&Y's success in knowledge management may be linked to their clear rewards given to employees who document how they have shared knowledge.

Collaboration and sharing of information are driven by processes, but can be enabled by technology. AutoSTEP has found that well-defined processes are key to making data exchange and sharing work. Tools, technology and standards support and enable the exchange process

The real showstoppers relate to cultural issues and perceptions. If people are unwilling to contribute to or use knowledge in the Knowledge Base, knowledge management *will* fail. Conventional processes tied to older and slower ways of doing business are entrenched in most organizations. Faster paced, collaborative work environments apply stress to conventional organizational structures. Individuals working on cross-functional project teams "break the rules" when they collaborate and appear to circumvent traditional organizational structures and line management. Organizations must adapt and reshape themselves to the new environments for knowledge management to be effectively implemented.

Telephone Survey Instrument

1. What is the scope of the system?
2. What are the goals of the system?
3. What methodologies were used in the implementation?
4. What standards were used in architecting the system?
5. What products were used in implementing the system?
6. What resources were required to develop the system?
7. What resources were required to operate the system?
8. What worked well in the implementation?
9. What unexpected events occurred in the implementation?
10. Who are your users? How many are there and what are their needs?
11. How do you communicate with your users?
12. What collaboration methods and tools are used?
13. What is done to enhance information?
14. What is done to keep information current?
15. What data formats do you support?

Figure A-2. Telephone Survey Questions

B Standards

This appendix provides more detailed descriptions of each of the standards referenced in Sections 4 and 5.

The standards included in JPL's Knowledge Management Architecture document are not meant be an exhaustive list of standards applicable to the implementation of Knowledge Management. Rather, they consist of the essential and emerging standards to establish a knowledge management framework at JPL. In the course of implementing knowledge management, additional standards will be specified in *A Knowledge Management Implementation Plan for JPL* related to specific disciplines and specific applications.

Also, it is important to note that "not all standards are created equal". Standards have life cycles and different rates of maturity. Therefore, it is important that standards be reviewed at regular intervals and evaluated in terms of their relevance to the current working environment. In addition, some of these standards are still in their infancy and a limited set of vendor technologies may be available to support these at this time. They are however, designed with an open architecture in mind, and the JPL knowledge management effort should closely monitor their progress in adoption by both Standards organizations and the commercial sector.

The standards described in this Appendix are noted below:

DMA	ODMA	WebDAV	WfMC
DSSSL	HTML	IGES	JPEG
MPEG-1	PDF	PNG	SGML
ODMG	RDA	SQL	SQL-3
ActiveX	CORBA	CORBA Services	UML
X.500	C	C++	Java
PCTE	Internet Official Protocol	IP	ISDN
TCP	Videoconferencing	LDAP	ICAP
vCalendar	MDIS	OIM	Metadata Coalition
XIF	H.323	T.120	PICS
XML	DOM	SAX 1.0	USMARC
Dublin Core	STEP	RDF	

B.1 DMA

Standard Name: Document Management Alliance (DMA)
Designation: DMA 1.0 Specification
Related Designations: ODMA, WebDAV
API: DMA API
Sources of Information: http://www.aiim.org/dma/dma_exec_overview.html

The DMA is both an organization and a specification. It is an organization of 50+ commercial and government organizations, which includes users of Data Management Systems (DMS), vendors, systems integrators, consultants, and industry analysts. The DMA organization is a task force of the Association for Information and Image Management (AIIM). One of AIIM's charters is to help establish consortiums that help set industry direction. DMA is dedicated to defining interoperability standards for document management products.

The DMA Specification is an interface standard that enables DMSs from different vendors to interoperate. After more than two years of development and testing, the 1.0 Specification was approved by DMA in December, 1997.

The DMA standard provides a rich set of capabilities, including

- A mechanism for automatically locating repositories
- The capability to map common attributes across repositories, even when those attributes have different names on different repositories
- Support for versioning
- Support for folders
- The ability to browse across DMSs using Explorer or internet browsers
- The ability to manage multiple renditions of a document
- Automatic discovery of document classes, properties, and search operators
- The ability to search across multiple repositories simultaneously, and merge the search results
- Full international support (incl. wide characters) and more method for delivery of documents to clients

The standard covers both client and server interfaces, enabling the broadest possible level of interoperability. Figure A-1 shows how DMA components fit together at the end user layer (client application/browser) and server.

The Document Management Alliance (DMA) 1.0 Specification creates the industry's first standard enabling document management systems from different vendors to interoperate. The enormous growth in the use of Electronic Document Management Systems (DMSs) has highlighted the need for interoperability among systems from different vendors. Without this interoperability, organizations end up creating “islands” of information and may not have the flexibility to choose the best products for their domain-specific needs.

Electronic Document Management Systems or DMSs, are commercial off-the-shelf or custom software packages that are cousins to Data Base Management Systems (DBMS). Whereas DBMSs store “structured data”—short records such as a name, address, account number, and social security number—DMSs store, retrieve, and manage “unstructured data” such as files, text, spreadsheets, images, sound clips, multi-media, and compound documents. The majority of these deployments to date have been departmental (single system) and they haven’t had to with enterprise level problems across multiple systems from multiple vendors. This is also the case at JPL, where we have a mix of commercial, custom developed, and mixed DMS systems in use, often implemented at the project level.

Because applications vary from institution to institution, and department to department, a one-size-fits all strategy for document management is not practical. Organizations need the flexibility pick the best-of-breed DMSs to fit their needs and the ability for systems from different vendors to interoperate. Without this interoperability, companies are forced to purchase from a single vendor, or—if they choose to purchase from multiple vendors—are forced to implement “islands” of information which are isolated from each other. The lack of interoperability also has the effect of increasing costs, by decreasing competition. And it makes integrating new technology with “legacy” systems much more difficult and expensive.

The internet explosion has contributed significantly to the rapid deployment of DMSs. First, internet technology makes it more economical to deploy DMSs by using browsers instead of expensive proprietary client software for every desktop. Today, many document management vendors offer a combined solution with internet connectivity through a Web gateway.

As corporations put internal and external documents on their intranets and Web sites, it often spotlights the lack of processes associated with storing, securing, finding, changing, and obsoleting corporate documentation. DMSs solve these problems, but when a company purchases systems from different vendors, coordinating such things as security and version control across multiple, incompatible systems becomes an integration challenge. Putting documents up on the Web instantly magnifies all of them and also all of the problems with managing corporate documentation.

DMA members are starting to implementing the standard into their products, and DMA-compliant products are beginning to on the market (as shown in Figure B-1).

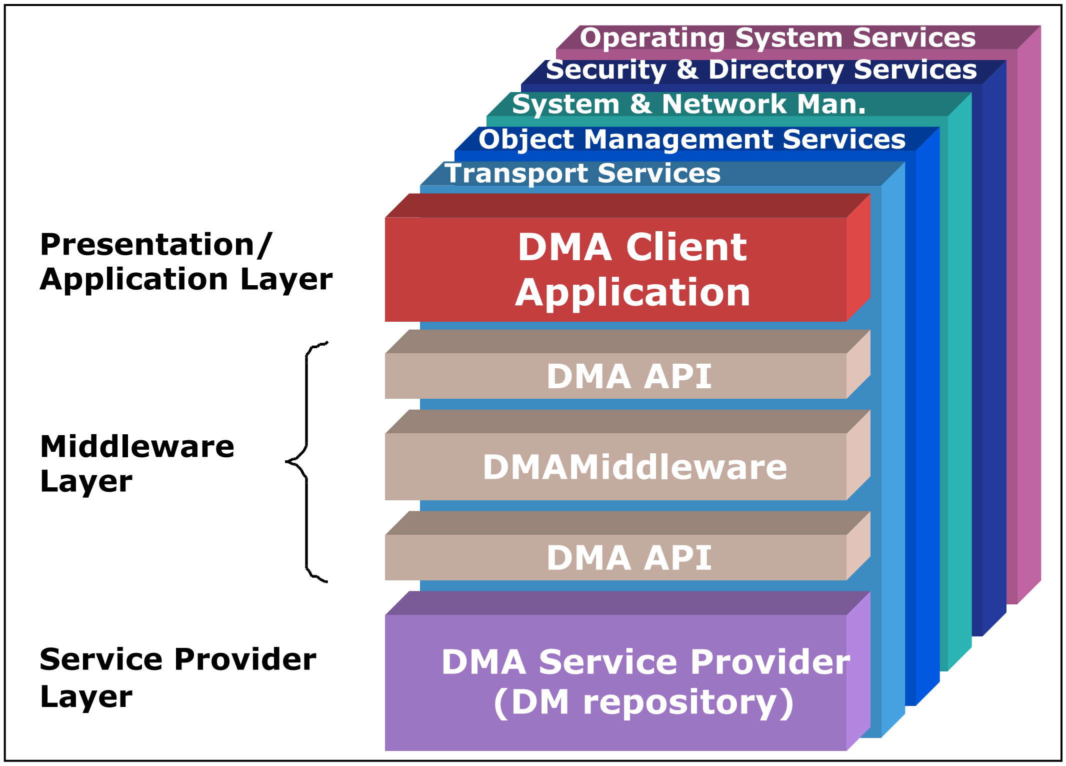


Figure B-1. DMA Components Fitting Together at the End User Layer

The DMA effort is ambitious and sophisticated, because it means that any conforming client, including Web browser clients, interact with any conforming DMS without having to know in advance the specific commands and characteristics of each DMS. It enables a client to use its own Graphical User Interface (GUI) and command set to store and retrieve objects from different-vendor DMSs and to discover DMS characteristics when a request is first sent. The features identified above and many more are part of the DMA 1.0 Specification, which was formally approved in December 1997. The priorities for the DMA 2.0 and later levels of the specification will be determined by continued user feedback.

Applicability to JPL: This standard is one to watch closely for product selection within the Document and Data Management Service. JPL already has several document management in use in house, both commercial and custom. Document exchange requirements with partners and sponsors will increase the need for interoperable DMSs. The Xerox DocuShare system in use by several projects is tracking the DMA specification, and a version due out in 1999 will provide DMA interoperability.

B.2 ODMA

Standard Name:	Open Document Management API (ODMA)
Designation:	ODMA
Related Designations:	DMA, WebDAV
API:	ODMA API
Sources of Information:	http://www.aiim.org/dma/dma_exec_overview.html

A few years ago, document management standards efforts were started at two levels. One was focused on a simple application programming interface (API) to let any kind of client interact with a that also implemented the API, for the purpose of storing and retrieving files. Desktop applications like word processing and spreadsheet packages are on those clients and must interact with the Data Management System (DMS) to store and retrieve the files created with those packages. The DMS replaces the Windows file system/director. With this standard, the client must know the specific design, construction, capabilities, etc., of the DMS in order to use it, including its proprietary document structuring, indexing, and query facility. Because all this knowledge is inside the client, the API itself is simple and inexpensive, yet so valuable because it makes the power of a proprietary DMS available to a wide range of desktop applications.

This API standard (Figure B-2), called ODMA (for Open Document Management API) has been built into many different kinds of clients, and is used widely today. It can be viewed as a many-to-one standard, many different clients to interact with each proprietary DMS in each DMS's own proprietary way. Because each client must be intimately knowledgeable in advance of each DMS with which it interact, it does a portion of the interoperability job needed by our example, but falls far short of the whole job.

In parallel, a second, more ambitious standards effort was launched to create interoperability across the different proprietary DMSs regardless of the platforms on which they reside and regardless the networks in which they exist, and without requiring clients to have advance intimate DMS knowledge—DMA. DMA can be combined with the ODMA standard for universal client access, and adds what is needed for completely vendor-independent cross-repository interoperability. The DMA specification provides many-to-many interoperability. That's many clients to many DMSs, of platforms and networks. In addition, because it accommodates international multi-language conventions, it's even language-independent.

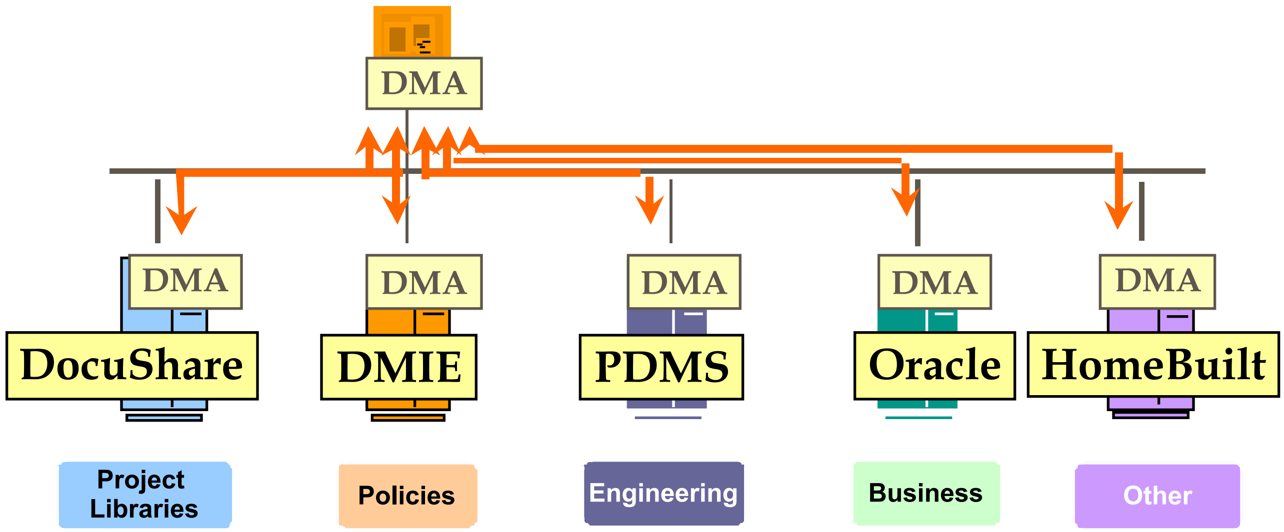


Figure B-2. ODMA Application Programming Interfaces

Applicability to JPL: This standard is supported by many of JPL standard desktop applications, including MS Office and Visio. In addition, DocuShare V2.0 due out by March 1999 will allow Windows desktops direct saving of files from ODMA compliant applications directly into DocuShare repositories via the ODMA API.

B.3 WebDAV

Standard Name:	Web Distributed Authoring and Versioning
Designation:	WebDAV
Related Designations:	DMA, ODMA
API:	WebDAV API
Sources of Information:	http://www.aiim.org/dma/dma_exec_overview.html

Another standards effort currently in process is called Web Distributed Authoring and Versioning (WebDAV) (Figure B-3). WebDAV is working to define extensions to the internet HTTP protocol which would Web pages created with one Web authoring tool, such as Microsoft's FrontPage, to be revised using a different Web authoring tool, for example Netscape's Navigator Gold. Because the WebDAV group has proposed including features such as "checking out" a Web page, or tracking versions of Web pages (which are features commonly found in document management systems), there has been some confusion about WebDAV possibly competing or conflicting with DMA.

In reality, WebDAV and DMA are extremely complementary. DMA creates interoperability of Web servers with a variety of document repositories. WebDAV creates interoperability of the tools used to author and revise Web pages. Both have the effect of increasing openness and interoperability for Web-based applications. In fact, the WebDAV and DMA organizations are actually collaborating to help insure that these standards align with one another, so that users can in fact use them in conjunction.

Applicability to JPL: This is a standard to watch and consider in product evaluation and selection for authoring components of the Web Site Management Service.

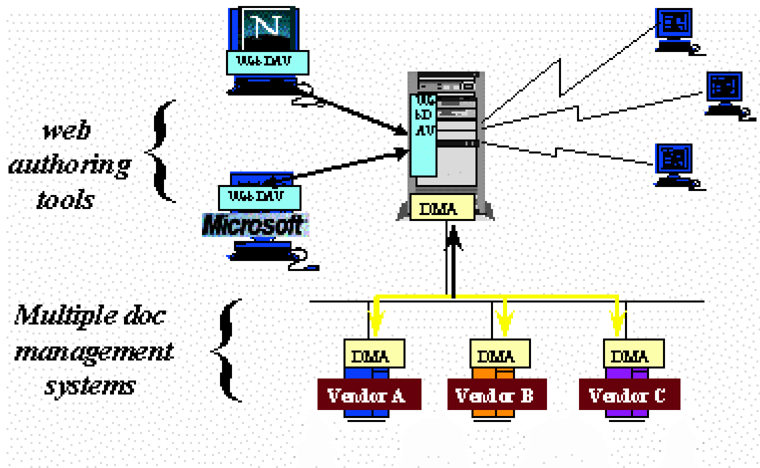


Figure B-3. WebDAV Distributed System

B.4 WfMC

Standard Name:	Workflow Management Coalition (WfMC)
Designation:	WfMC
Related Designations:	DMA
API:	WfMC API
Sources of Information:	http://www.aiim.org/dma/dma_exec_overview.html

A related standards effort to the DMA initiative is known as the Workflow Management Coalition. Just as DMA has created an interoperability standard for document management systems, WfMC has defined an interoperability standard for workflow software products. Workflow tools—which allow users to automate and control their work processes—are often used in conjunction with document management systems. WfMC allows things such as a workflow program to be exported from one tool and imported into another, or a work item to be shared by two different workflow software packages.

Because document management systems and workflow are so often used together, the WfMC initiative is highly complementary to DMA, and can be used in conjunction with DMA. The DMA and WfMC organizations demonstrated these standards working together at major industry trade shows in 1998, including the AIIM conference in Anaheim in May 1998.

Applicability to JPL: Workflow interoperability standards are worthy goals, but this standard may be too difficult to implement for a large number of vendors. JPL should watch this and alternate data exchange methods such as SAX for market adoption.

A.5 DSSSL4

Standard Name:	Document Style Semantics and Specification Language (DSSSL)
Designation:	ISO/IEC 10179:1996
Related Designations:	None
API:	NA
Sources of Information:	Contact ISO/IEC for the specification · http://www.jclark.com/dsssl/ · http://navysgml.dt.navy.mil/dsssl.html · ftp://infosrv1.ctd.ornl.gov/pub/sgml/WG8/DSSSL

The Document Style Semantics and Specification Language (DSSSL) provides two separate and independent processes: the Standard Generalized Markup Language (SGML) Tree Transformation Process (STTP) and the SGML Tree Formatting Process (STFP). STTP is a structural tree transformation process for converting an SGML (refer to D.6.11 for information on SGML) document structure specified by one Document Type Definition (DTD) to a different document structure specified by another DTD by mapping the elements of one DTD structural tree to the elements of another DTD structural tree.

STFP is a process for applying formatting information to an SGML document structural tree by specifying formatting parameters for each tagged element of the document. The formatting application of DSSSL also provides the page layout specifications for page-oriented output applications. Each application of DSSSL formatting is specific to the output device being used. A separate DSSSL specification is required for each different output device and each different output format. The DSSSL output imaging model is PostScript Level 2 and the font model is specified in ISO/IEC 9541.1-3.

[4] This standard definition is taken from [61]. This also applies to the descriptions for HTML, IGES, JPEG, MPEG-1, PDF, PNG, SGML, ODMG, RDA, SQL, SQL-3, ActiveX, CORBA, CORBA Services, UML, X.500, C, C++, JAVA, DOM, USMARC, Videoconferencing, ISDN, TCP, IP, Internet Office Protocol, and PCTE.

B.6 HTML

Standard Name: HyperText Markup Language (HTML)
Designation: HTML 2.0
Related Designations: None
API: NA
Sources of Information: Contact IETF and W3C for the specification

- http://www.cwru.edu/help/introHTML/toc.html
- http://www.rochester.edu/Help/htmlinfo.htm
- http://union.ncsa.uiuc.edu/HyperNews/get/www/html/lang.html

The HyperText Markup Language (HTML) is a Standard Generalized Markup Language (SGML) Document Type Definition (DTD) for creating hypertext documents that are portable across heterogeneous platforms. HTML is a data format used to create hypertext documents from unformatted text that includes embedded tags that define the structure of the document. Tables, forms, and links to graphics and other Web pages can be added. Refer to D.6.11 for information on SGML and DTDs.

The HTML DTD defines the HTML syntax in terms of SGML. When an HTML document is displayed by non-HTML-enabled word processing software, the tags appear as character strings embedded within the unformatted text. When displayed by HTML-enabled software, the text appears formatted according to the convention indicated by the tags, which are not displayed.

The HTML 2.0 specification is maintained by the Internet Engineering Task Force (IETF) as Request for Comment (RFC) 1866. This specification also defines HTML as an Internet Media Type [IMEDIA] and MIME Content Type [MIME] called “text/html”. HTML 2.0 contains SGML Document Access (SDA) fixed attributes for document transformation to the International Committee for Accessible Document Design (ICADD) DTD. ICADD applications enable vision-impaired persons to access structured documents through Braille, large print format, and voice synthesis as defined in ISO 12083:1993.

The HTML standard allows Web browsers to include private extensions for features not supported by the HTML standard and most Web browsers include their own set of private extensions. Incorporating these features can enhance the functionality and appearance of an HTML document, but can also make that part of the document unreadable by other browsers that do not support that feature.

B.7 IGES

Standard Name: Initial Graphics Exchange Specification (IGES)
Designation: FIPS 177-1
Related Designations: ANSI/USPRO/IPO-100-1993 Ver 5.2
API: NA
Sources of Information: Contact NIST for a copy of the FIPS and ANSI for the specification

- http://www.scra.org/uspro//stds/wh-iges.html
- http://www.scra.org/uspro//question/whtsiges.html
- http://speckle.ncsl.nist.gov/~jacki/igests.htm
- U.S. Product Data Association, Suite 200, 2722 Merrilee Drive, Fairfax, VA 22031 Telephone: 703-698-9606

Initial Graphics Exchange Specification (IGES) is a set of protocols that specifies a mechanism for the digital exchange of database information among computer-aided systems. It provides a data format for describing product design and manufacturing information that has been created and stored in computer readable form. IGES information, including drawings, three-dimensional wire-frame models, and surface models, is intended for human interpretation at the receiving site. The IGES format is designed to be independent of all CAD/CAM systems.

IGES specifies the data required to describe and communicate the essential engineering characteristics of physical objects such as manufactured products. Such products are described in terms of their physical shape, dimensions, and information which further describes or explains the product. The processes that generate or utilize the product definition data typically include design, engineering analysis, production planning, fabrication, material handling, assembly, inspection, marketing, and field service.

A format to allow the exchange of a product definition between CAD/CAM systems must, as a minimum, support the communication of geometric data, annotation, and organization of the data. The file format defined by IGES treats the product definition as a file of entities. Each entity is represented in an application-independent format, to and from which native representation of a specific CAD/CAM system can be mapped.

B.8 JPEG

Standard Name: Joint Photographic Experts Group (JPEG)
Designation: ISO/IEC IS 10918.1-2
Related Designations: None
API: NA
Sources of Information. Contact ISO/IEC for the specification

- http://www-cse.ucsd.edu/users/jhan/hjpegtut.html
- http://www-personal.umich.edu/~jweise/quality/JPEG.HTML

The Joint Photographic Experts Group (JPEG) standard is a family of compression algorithms for gray-scale and color still images of natural, real-world scenes with a pixel depth of 6 to 24 bits. JPEG works best on photographs, naturalistic artwork, complex computer drawn images (ray-trace scenes), semi-realistic artwork (fantasy drawings), and similar images. JPEG does not work well with text, simple cartoons, simple artwork, and similar images. JPEG is stored in full 24 bits per pixel color (16 million colors). File formats that encode or decode JPEG images are

- JPEG File Interchange Format (JFIF)
- Tag Image File Format (TIFF) version 6.0
- Encapsulated PostScript (EPS)

JPEG defines a baseline lossy algorithm, plus optional extensions for progressive and hierarchical coding. The JPEG lossy algorithm discards information on slight changes in color (chromanance) that the human eye cannot easily detect but retains changes in light intensity (luminance). Repeated compression/decompression cycles of the same image can result in a visually-degraded image.

JPEG has a selectable quality, or Q, factor that determines the amount of compression, from 0 to 100, that will be applied to the image. JPEG can obtain 10:1 to 20:1 compression without visible lost of image quality and 30:1 to 50:1 with some visible lost of image quality. The threshold for quality lost for gray-scale images is about 5:1. The amount of compression that can be applied before changes in the image are noticeable depends greatly on the resolution, size, detail, and color of the image.

JPEG has two coding forms for image output: Huffman coding and arithmetic coding. The choice of image output coding has no affect on image quality. However, arithmetic coding usually has better compression by 5 to 10 percent.

Sequential encoding involves encoding an image in a single left-to-right and top-to-bottom scan. The most widely used mode of JPEG is the baseline mode which is a subset of sequential encoding based on the Discrete Cosine Transform (DCT) compression algorithm.

The progressive mode extension is for real-time transmission of images. It allows the DCT coefficients to be sent piecemeal in multiple scans of the image. With each scan, the decoder produces a higher-quality rendition of the image. This mode allows a low-quality preview image to be sent very quickly, then refined to produce a higher-quality image. This mode is useful only with fast decoders or slow transmission lines.

The hierarchical mode represents an image at multiple resolutions. For example, 512x512, 1024x1024, and 2048x2048 versions of the image can be provided. The higher-resolution images are coded as differences from the next smaller image. Hierarchical mode is not widely supported.

B.9 MPEG-1

Standard Name:	Motion Picture Experts Group (MPEG)
Designation:	ISO/IEC IS 11172.1-4
Related Designations:	None
API:	NA
Sources of Information:	Contact ISO/IEC for the specification

- http://wheat.symgrp.com/symgrp/datx/mpegwp.html
- http://www.optivision.com/compress/technica/wpaps2.html
- http://www.optivision.com/compress/technica/wpaps3.html
- http://www.cdrevolution.com/text/mpeginfo.htm

MPEG-1 defines a set of international standards for the compression and decompression of full-motion digital video signals at compression ratios up to 200:1. MPEG-1 provides video quality approaching that of VHS videos.

MPEG-1 is for use with computers, games, and set-top boxes providing high-quality, full-motion video for multimedia applications. MPEG-1 format is used for Web pages, PC-based training videos, presentations, kiosk programs, and product demonstrations.

MPEG-1 specifies a video resolution of 352x240 pixels compressed at 30 frames per second (fps) at a bandwidth of 150 kilobits per second (kbps). Sampling rates as high as 4095 x 4095 x 60 are allowed, but rarely used. There are three ways to encode MPEG-1: with a multiple processor system; with a real-time capture encoding solution; or with hardware and software combinations that produce MPEG-1 in two separate steps. A video board with an MPEG decoder chip is needed to view an MPEG video.

An MPEG-1 stream has two layers: the system layer containing timing and other information needed to demultiplex the audio and video streams and to synchronize audio and video during playback, and the compression layer which includes the compressed audio and video data. MPEG compression uses both intraframe and interframe encoding which takes advantage of temporal redundancy¾ the fact that video is usually made up of many successive frames that contain large areas that do not change from frame to frame¾ by noting differences between consecutive frames.

MPEG has three frame types: intra (I), predicted (P), and bidirectional (B). An I frame is encoded as a still frame, with no information about previous or future frames. A P frame is predicted from the most recently encoded I or P frame and uses motion compensation to provide more compression than with I frames. If there is no good match to the previous I or P frame, the P frame is encoded as an I frame. A B frame is predicted from the closest two I or P frames where one of the closest frames is in the past and the other is in the future. These two closest frames can be considered separately or averaged to find a good match. B frames provide the most compression. Bidirectional prediction also decreases the effect of noise by averaging two frames. If no good match is found, the B frame is encoded as an I frame.

MPEG provides a timing mechanism that ensures synchronization of audio and video signals. Two parameters are used by the decoder; the system clock reference (SCR) and the presentation timestamp (PTS).

An SCR is a snapshot of the encoder system clock. The SCRs used by the audio and video decoder must have approximately the same value to keep the audio and video synchronized. The video and audio decoders update their internal clocks using the SCR value sent by the system decoder.

PTSs are samples of the encoder system clock that are associated with some video or audio presentation units. A presentation unit is a decoded video picture or a decoded audio time sequence. The PTS represents the time that the video picture is to be displayed or the starting playback time for the audio time sequence. If the PTS is earlier than the current SCR, the video decoder discards the picture. If the PTS is later than the current SCR, the video decoder repeats the display of the picture.

B.10 PDF

Standard Name: Portable Document Format (PDF)
Designation: Adobe PDF
Related Designations: PostScript (PS)
API: NA
Sources of Information: Information on PDF can be obtained from

- http://www.adobe.com/
- http://www.vertec.com/info/total.htm
- Adobe Systems Incorporated, 345 Park Avenue, San Jose, California 95110-2704 Telephone: 408-536-6000; Fax: 408-537-6000

PDF is a final form, proprietary, lossless bitmap document file format specification owned by Adobe Systems. It is an open and extensible file format allowing implementation of third-party software plug-ins. The specification for the standard is published in *Portable Document Reference Manual* (Tim Bienz and Richard Cohn, Addison-Wesley, 1993). Changes to the standard are made by Adobe's PDF Language Committee and are published, among other places, on the Adobe Web site (http://www.adobe.com/).

PDF takes its imaging model and device independence from PostScript; both paint an image one dot, or pixel, at a time. Converting a document to PDF is similar to printing a document to a PostScript printer. Once a document is in PDF, it can be read or viewed, but not altered, on any computer platform that supports PDF.

Although PDF can be used to store and display any document, it is specifically beneficial when maintaining the exact appearance or design of the document in an uneditable format is important. Typical examples are financial reports, legal documents, technical reports, handbooks, and catalogs. PDF is also suited for storing electronic images of paper documents when there is no original electronic file or the electronic file is in a format that is no longer readable.

Password protection is not a part of the PDF standard, but PDF documents can be password protected and the password can be encrypted. Encryption and decryption are supported in PDF with independent plug-in applications.

B.11 PNG

Standard Name: Portable Network Graphic (PNG)
Designation: W3C-REC-PNG-961001
Related Designations: None
Sources of Information: Contact W3C for the specification

- http://www.w3.org/pub/WWW/Graphics/PNG/

PNG is an editable, nonproprietary, lossless bitmap graphics file format standard maintained by the World Wide Web Consortium (W3C) that has been approved as a W3C Standard: W3C-REC-PNG-961001. It was developed in conjunction with CompuServe as a replacement for GIF. PNG is also well suited for true color ray-traced images, which do not compress well using Joint Photographic Experts Group (JPEG) compression.

The maximum image size allowed is 2^{32}B1 by 2^{32}B1 pixels. An image can have an indexed color pixel depth up to eight bits using up to a 256-color palette lookup table indexed with the RGB color model, an RGB true color pixel depth up to 48 bits (up to 16 bits for each of the R, G, and B samples), or a grayscale pixel depth up to 16 bits. An optional alpha channel is available to provide 8-bit transparency data on a per-pixel basis for true color and grayscale images. The allowed color, grayscale, and alpha channel combinations are provided in Table B-1.

Table B-1. Allowed Color and Grayscale Combinations

Allowed Bit Depths/Sample	Sample Description	Alpha Channel
1, 2, 4, 8, and 16	Each grayscale pixel sample	No
8 and 16	Each true color pixel sample for R, G, and B	No
1, 2, 4, and 8	Each indexed-color pixel sample	No
8 and 16	Each grayscale pixel sample	Yes
8 and 16	Each true color pixel sample for R, G, and B	Yes

The deflate/inflate compression algorithm with an LZ77-type 32-kbit sliding window is the only one currently allowed. Compressed data streams are stored in the zlib format specified in IETF RFC-1950 using the deflate compression setting.

The scan lines in a PNG image can be stored in consecutive order, beginning with the first row at the top. PNG also supports displaying progressive or interlaced images using the Adam7 two-dimensional, seven-pass interlacing sequence.

PNG provides optional image gamma information for the device on which the image was created. This information can be used by the viewing device to correct for brightness and color differences between the two devices. Additional color correction enhancements can be achieved with an optional chromaticity setting. This chromaticity information provides more accurate grayscale conversion, provides better RGB color fidelity on monitors, provides better color fidelity when converting from RGB to cyan-yellow-magenta-black (CYMK) for printing, allows calculating an optimal color palette, and allows transferring the image to a color management system.

PNG also provides file integrity checking and error detection and correction. A file must contain specified information blocks (PNG calls them chunks) in a set sequence beginning with an eight-byte file signature that identifies the file as a PNG file, followed by a header block and ending with an end block. Naming conventions for all intermediate blocks are also specified. A four-byte cyclic redundancy check (CRC) algorithm, as defined in ISO/IEC 3309, is used in each data block to ensure accurate transmission of the data.

Other PNG features support overlaying ASCII text on the graphic, stopping display of the graphic for user input, incorporating application-specific information to enhance the image display, and incorporating undisplayed, human-readable ASCII text in the data stream.

PNG does not support multiple images in a single file, although this is not necessarily a limitation since most readers do not support multiple images per file or display only the first image in the file.

B.12 SGML

Standard Name:	Standard Generalized Markup Language (SGML)
Designation:	FIPS 152
Related Designations:	ANSI/ISO 8879:1986
API:	NA
Sources of Information:	Contact NIST for a copy of the FIPS and ANSI/ISO for the specification · http://www.sgmlopen.org/ · http://www.sil.org/sgml/sgml.html · http://www.arbortext.com/wp.html

The Standard Generalized Markup Language (SGML) is a meta-language that specifies the methodology used to encode or tag the structural elements of a document. SGML does not specify the actual tags that must be used, although certain tag sets have become de facto standards by general usage. SGML documents can be seamlessly exchanged between heterogeneous systems.

A document can be envisioned in three layers: structure, content, and style. Structure is the organization and hierarchy of the document, content is the actual elements (text, tables, graphics) of the document, and style is the appearance (page layout, type font, point size, etc.) of the document. SGML only addresses the relationship between the structure and content of the document.

Content is the actual textual information in the document. Content includes the text of titles, sections, lists, and tables. Content also includes graphics and audio.

Structural tagging, also called descriptive or generic markup, defines the purpose of the text elements in a document, rather than the physical appearance of the element on the page. Structural tags are inserted immediately before and after each element. These tags identify elements within the document structure¾a chapter, section, list, etc.¾using notations that describe what each element is, not how it appears. By separating the style of a document from its structure, structural tagging allows for multiple presentations of the same information from the same source file. For example, printing on paper, displaying on-line, or copying onto a CD-ROM.

SGML provides a methodology for setting up hierarchical models, called document type definitions (DTDs), for each type of document. A DTD specifies a hierarchical framework for the elements of the document (such as chapters, headings, and paragraphs). The DTD also specifies the relationships between structural elements; for example, "a paragraph must be the first element after a paragraph heading" or "each list must contain at least two items."

A different DTD is usually created for each type of document: reports, design specifications, technical manuals, catalogs, letters, and memos. General industry-standard DTDs are available. Two of these are HyperText Markup Language (HTML) DTD and Electronic Manuscript Preparation and Markup (EMPM) DTD as specified in ANSI/NISO Z39.59B1987.

B.13 ODMG

Standard Name:	Object Data Management Group (ODMG) Standard
Designation:	ODMG 1.2 (1993)
Related Designations:	Standard references OMG specifications, and is almost compatible with SQL object query language.
API:	Yes
Sources of Information:	For copies of the specification see computer bookstores · http://www.odmg.org/odmg-93.html · Morgan Kaufmann Publishers, 340 Pine Street, Sixth Floor, San Francisco, CA 94104; Sales: 800-745-7323

Standard for object-oriented databases such as that provided by the following vendors, GemStone Systems, IBEX Computing, O2 Technology, Object Design, Objectivity, POET Software, UniSQL, and Versant Object Technology. Version 2.0 of the standard is scheduled for release in 1997 and will contain a Java binding in addition to the current C++ and SmallTalk bindings.

B.14 RDA

Standard Name: Remote Database Access (RDA)
Designation: ISO/IEC 9579 (Parts 1B3)
Related Designations: OIW/X/Open RDA CAE Specification C307, Version 2
Note also that RDA uses a derivative of SQL-3 CLI as its API
API: Yes, as well as a format and protocol.
Sources of Information: Contact ISO/IEC for the specification
- http://www.livedata.com/index.html

RDA is a database interoperability middleware standard for client server database environments. Initially designed to run over OSI protocols, it has been adapted to also run over TCP/IP. RDA products allow a user to access multiple databases from different vendors using a common protocol and API. The RDA specification also defines an interface to ISO (transaction processing) two phase commit TP services in the case where updates to multiple remote databases need to be coordinated. RDA has not been very successful in the commercial market place due to lack of support from the major DBMS vendors and the availability of other gateway and proprietary solutions. If this standard ever gains vendor support, it will save the user community money through elimination of the fees currently paid for client database drivers and proprietary database middleware software. The current primary users of this standard are the Air Force and a consortium of utility companies.

- ISO/IEC 9579-2 SQL Specialization¾ http://www.iso.ch/cate/d17345.html
- ISO/IEC 9579-1 Generic Model, Service and Protocol¾ http://www.iso.ch/cate/d17344.html
- ISO/IEC 9579-3 SQL Specialization Protocol-Implementation Conformance Statement (PICS) proforma¾ http://www.iso.ch/cate/d24405.html

Also a copy of the Open Systems Environment Implementers Workshop (OIW) Draft Implementor Agreement/X/Open RDA CAE Specification C307 is available on the internet at nemo.ncsl.nist.gov/oiw/rdasig/ti-rda4.rtf.

NIST has done a lot of prototyping with RDA and has published the following RDA/SQL Toolkit and ODBC Programmer's Toolkit on the internet at alpha.ncsl.nist.gov/~anonymous/RDATOOLKIT.html.

B.15 SQL

Standard Name:	Database Language SQL
Designation:	FIPS 127-2
Related Designations:	ISO/IEC IEC 9075:1992, ANSI X3.135-1992, SQL-92
API:	NA. SQL serves the same purpose as an API; SQL provides the programmer with specialized programming language statements rather then defined procedure calls.
Sources of Information:	Contact NIST for a copy of the FIPS, ANSI and ISO/IEC for the specification · http://www.nist.gov/itl/div897/pubs/fip127-2.htm · Jim Melton and Alan R. Simon, *Understanding the New SQL: A Complete Guide*, Morgan-Kaufmann Publishers, San Mateo, CA 94403, October 1992

SQL is a data sub-language used to access relational databases. It is widely supported by the major database server vendors.

B.16 SQL-3 Call Level Interface

Standard Name: SQL-3 Call Level Interface
Designation: SQL-3 Part 3, also approved addendum to SQL-92
Related Designations: ODBC 3.0 (1996)
API: Yes
Sources of Information: Contact ANSI for the specification

- http://www.openlink.co.uk
- http://www.visigenic.com
- http://www.ibi.com/booksibi/summary.html
- http://www.intersolv.com/products/add_home.htm
- http://www.synergex.com/odbc/odbc_man.htm
- http://www.UNISYS.com/softstor/2200/22uni.html (for Unisys 2200 RDMS ODBC access)
- http://www.NCR.com/product/teradata/prodinfo/odbcnew.htm
- http://www.microsoft.com/odbc regarding ODBC 3.0
- http://www.pcinews.com/business/pci/sun/features/odbc.overview.html

An extension to the SQL standard to cover such functions as database connection setup and tear down, database environment detection, and the encapsulation of SQL statements in a procedure call structure. The database may be either remote or local. The standard is widely supported. Drivers or gateway products must be procured to access specific target databases. The Windows implementation of this standard is called Open Database Connectivity (ODBC) and is now being ported to UNIX. A Java binding was also recently announced which is called JDBC. Procurers must be aware of the following variable capabilities: single tier versus two tier drivers; 16 bit versus 32 bit implementations; version level; functionality supported level¾ core, level 1, and level 2; and level of SQL grammar support.

B.17 ActiveX

Standard Name:	ActiveX
Designation:	ActiveX
Related Designations:	Object Linked Embedding (OLE)
API:	Yes, API's are associated with ActiveX and the underlying DCOM, however, ActiveX is more than a set of API's. An ActiveX control is a component that meets the ActiveX definition and interacts with its environment using the COM. There are several ways to write ActiveX controls see "Sources of Information," but a library such as Microsoft's Foundation Class is usually involved. To use an ActiveX control written by someone else in your Web page involves setting parameters associated with the embedded object tag.
Sources of Information:	http://www.activex.org · http://www.microsoft.com/intdev/controls/controls.htm#def for a definition and on how to start writing ActiveX controls · North, "Database Programming with OLE and ActiveX." *DBMS* (November): 87-92 (1996)

ActiveX is a significant object-based standard for use primarily on Windows platforms. ActiveX has been ported to some UNIX environments, but is not well established on UNIX. ActiveX is a slimmed down subset of OLE optimized for the internet. The term ActiveX is usually used to refer to ActiveX controls, which are similar to OLE controls. The term ActiveX may also be used to refer to ActiveX automation interfaces, or ActiveX data objects. Microsoft defines ActiveX as "a set of technologies that enables software components to interact with one another in a networked environment, regardless of the language in which they were created." ActiveX used to belong only to Microsoft, but in the fall of 1996 Microsoft ceded ownership of the standard to "The Open Group" standards consortium. ActiveX is built on the Component Object Model (COM). ActiveX is a competitor to the Object Management Group (OMG) Common Object Requester Broker Architecture (CORBA) set of standards and is only recommended in those situations where CORBA is not practical. Currently the industry is divided between Microsoft and The Open Group with ActiveX, OLE, Component Object Model (COM), and Distributed COM (DCOM) on one side; and Netscape, IBM, OMG and UNIX vendors, with CORBA, Live Objects, DSOM on the other side. Plans and standards are already in place for coexistence of these two sets of standards. OMG objects can encapsulate OLE objects and vice versa. It remains to be seen as to whether one set of standards or the other will prevail. In the interim it is clear that both will exist.

B.18 Common Object Request Broker Architecture

Standard Name:	Common Object Request Broker Architecture (CORBA) 2.0
Designation:	OMG document # 96-03-04
Related Designations:	See other object management architecture standards
API:	Yes and protocol definitions.
Sources of Information:	Contact OMG for the specification or http://www.omg.org

The CORBA 2.0 specification defines the protocols, such as internet Interoperability Protocol (IIOP), for communication between object request brokers (ORBs). ORBs serve as a clearing house for object-based communication. Netscape has embraced IIOP as the replacement for HTML in many internet applications. IBM's DSOM architecture is an implementation that uses CORBA.

B.19 CORBA Services

Standard Name:	CORBA Services
Designation:	OMG CORBA Services publication (see Table 5-2 for OMG document numbers)
Related Designations:	Common Object System Services (COSS)
API:	Yes
Sources of Information:	Contact OMG for the specification or http://www.omg.org

This standard defines sets of services for managing the object system, such as object creation and destruction, as well as less common services for events, naming, object queries, persistence, externalization, concurrency, transactions, relationships, security, and time.

B.20 Unified Modeling Language

Standard Name:	Unified Modeling Language (UML)
Designation:	UML version 1.0
Related Designations:	OMG in the future
API:	NA
Sources of Information:	http://www.rational.com

UML is a modeling standard similar to IDEF, but is based on methodologies that support object-oriented development. UML came from the merger of the Booch, Jacobson, Rumbaugh, object-oriented software engineering, and object modeling technique methodologies. The standard is to be submitted to the OMG in January 1997. Joint submitters include Rational, Microsoft, HP, MCI, Systemhouse, Oracle, TI, and Icon Computing. Once this standard becomes mature, this standard should be the choice for modeling efforts that will result in object-oriented (C++, Java, or SmallTalk) development.

B.21 X.500

Standard Name:	X.500 Directory Services
Designation:	ITU-T X.500
Related Designations:	ISO 9594
API:	Yes, the Directory Services API defines a standard directory service user agent interface to support application portability at the source code level. The API is contained in four documents: · IEEE 1224.2 Language Independent Specification (1993) · IEEE 1326.2 Language Independent Test Methods (1993) · IEEE 1327.2 C Language Binding (1993) · IEEE 1328.2 C Language Test Methods (1993) The C Language dependent API, IEEE 1327.2, is required. Users must require that vendors have demonstrated adherence of their API products to IEEE 1327.2 by testing according to IEEE 1328.2, C Language Test Methods. Although the Directory Services API is intended to provide access to X.500 functionality, its scope is not limited to X.500. It could be used to access other directory services as well.
Sources of Information:	See ITU-T X.500 and ISO 9594. Copies of the specifications can be obtained from ITU-T and ISO. The Validated Products List (VPL) can be accessed at ftp://speckle.ncsl.nist.gov/vpl/intro.htm

The OSI X.500 Directory Services (DS) standard provides a means of "looking up" certain object information. Examples of object information may include open system network addresses, e-mail addresses, and application addresses. The current trend for X.500 DS may include a mailhub backbone with access to on-line TCP applications, catalogs, identification cards, phone and fax service, voice mail, and API's to HTML ports.

The DS functional model main component is the Directory, which is composed of Directory Service Agents (DSAs). The DSAs communicate with other DSA elements by using the Directory Service Protocol (DSP) to provide services to users. The user, human or application program, requests service through a Directory User Agent (DUA). A DUA requests DSA services using the Directory Access Protocol (DAP). Both protocols conform to OSI definitions as specified by X.519.

A DUA interacts with the Directory by communicating with one or more DSAs and need not be bound to any particular DSA. DUA-DSA and DSA-DSA interaction occurs through requests and responses with request referrals. The three Directory topologies used to accomplish this interaction include chaining, multi-casting, and mixed mode hybrid (both).

The Directory Information Base (DIB) is a tree structured distributed hierarchy, referred to as the Directory Information Tree (DIT), used to store the object information. Directory retrieval patterns shall include object look-up, user-friendly naming, browsing, yellow pages (X.520), group(X.520), and authentication (X.509). The Directory shall support a number of generic applications such as inter-personal communications and OSI inter-system communications. A profile group may define its own DIT options which may adhere to DS standards or support custom requirements. X.500 Directory Service options may be defined by the CCITT specification, profile groups, or the software utility.

B.22 C

Standard Name:	C Programming Language
Designation:	FIPS 160, C
Related Designations:	ANSI/ISO 9899:1990, ANSI X3.159-1989, Tech. Corrigendum 1994,
API:	NA. C is a programming language. Supporting library functions constitute the available API.
Sources of Information:	Contact NIST for a copy of the FIPS and ANSI/ISO for the specification

- http://www.ansi.org
- http://www.iso.org
- ftp://speckle.ncsl.nist.gov/vpl/intro.htm for a copy of the VPL
- Brian Kernighan and Dennis Ritchie, *The C Programming Language, 2nd Edition*, Prentice Hall, 1988. The 2nd edition took into account changes to the language from the ANSI standards setting process.
- NIST Validated Products List (VPL)

C is a mid-level, 3rd generation procedural programming language. C is used both for operating system development as well as application development. C provides efficient, portable code.

B.23 C++

Standard Name:	C++
Designation:	ANSI/ISO X3J16 Committee Draft,
Related Designations:	AT&T C++
API:	Yes. Although C++ is a programming language, most C++ functionality is specified in callable routines from the *C++ Standard Library* in the draft specification.
Sources of Information:	Contact ANSI/ISO for the specification

- http://www.uni-duesseldorf.de/WWW/MathNat/Fokus/ub/cpp.html

- http://www.maths.warwick.ac.uk:80/c%2b%2b/pub/wp/

C++ is an object-oriented programming language written as an extension to the C programming language. C++ programs execute efficiently, but the language is relatively difficult to learn due to the large number of class libraries and APIs with which programmers must familiarize themselves. ANSI plans to finalize the specification during 1997.

B.24 Java

Standard Name:	Java
Designation:	Java 1.1
Related Designations:	Highly related but not identical standards include Java Virtual Machine Specification, Java Electronic Commerce Framework (JECF), JDBC, JMAPI, and Java Beans
API:	Yes. In addition there are a number of development environment extensions being pushed by the vendors.
Sources of Information:	kenpo.nmclites.edu/JavaDocs/apibook/javabook.htm for APIs · kenpo.nmclites.edu/JavaDocs/apibook/preface.htm for general description of Java · http://www.cs.auckland.ac.nz/~julian/java_resources.htm for links to other Java Web pages · Gosling, James; Joy, Bill; and Steele, Guy. *The Java Language Specification*. Addison-Wesley, Reading, Massachusetts, 1996, ISBN 0-201-63451-1 · Lindholm, Tim, and Yellin, Frank. *The Java Virtual Machine Specification*. Addison-Wesley, Reading, Massachusetts, 1996, ISBN 0-201-63452-X

Java is a network-oriented programming language invented by Sun Microsystems that is specifically designed for writing programs that can be safely downloaded to your computer through the internet and immediately run without fear of viruses or other harm to your computer or files. Using small Java programs (called “Applets”), Web pages can include functions such as animations, calculators, and other fancy tricks.

Sun has created a multiple-step process for developing new APIs or revising old ones. That process was one of the keys to Sun’s winning “official submitter” status for Java by the International Standards Organization.

A programming language for internet (World Wide Web) and intranet applications from Sun. Java was modeled after C++, and Java programs can be called from within HTML documents or launched stand alone. The first Web browsers to run Java applications were Sun’s HotJava and Netscape’s Navigator 2.0. Java was designed to run in small amounts of memory and provides its own memory management.

Java is an interpreted language that uses an intermediate language. The source code of a Java program is compiled into "byte code," which cannot be run by itself. The byte code must be converted into machine code at runtime. Upon finding a Java applet, the Web browser switches to its Java interpreter (Java Virtual Machine) which translates the byte code into machine code and runs it. This means Java programs are not dependent on any specific hardware and will run in any computer with the Java Virtual Machine. On the server side, Java can also be compiled into machine language for fastest performance, but they lose their hardware independence as a result.

Like other programming languages, Java is royalty free to developers for writing applications. However, the Java Virtual Machine, which executes Java applications, is licensed to the companies that incorporate it in their browsers and servers.

B.25 PCTE

Standard Name:	Portable Common Tools Environment (PCTE)
Designation:	ISO/IEC 13719:1995
Related Designations:	ISO/IEC 13719-1: Part 1 Abstract Specification ISO/IEC 13719-2: Part 2 C Programming Language Binding to PCTE, European Computer Manufacturers Association (ECMA)-149 and ECMA-158
API:	Yes
Sources of Information:	Contact ISO/IEC for the specification or http://gille.loria.fr:7000/pcte/

This standard is for the niche market of integrated software engineering environments exemplified by products such as Teamwork, TI Composer, and HP Workbench. Approximately thirty vendors sell either PCTE environments or tools. PCTE allows tools from different vendors to interoperate and share data in a common repository. The U.S. portion of this standard setting effort is being done under the umbrella of the OMG, because PCTE uses CORBA for enabling tool communication. A development environment that claims to be PCTE is more likely to be able to have third party add-in tools then one that does not, and therefore, PCTE environments and tools should be preferred in procurements over similar items that are not compliant.

B.26 Internet Official Protocol Standards

Standard Name: Internet Official Protocol Standards
Designation: RFC 1920
Related Designations: IAB Standard 1
API: NA
Sources of Information: http://ds.internic.net/ds/rfc-index.htm

RFC 1920 describes the state of standardization of protocols used in the internet as determined by the Internet Architecture Board (IAB). RFC 1920 is an Internet Standard with the designation of IAB Standard 1. The IAB recommends that the standards process be used in the evolution of the protocol suite to maximize interoperability and to prevent incompatible protocol requirements from arising. Few protocols are required to be implemented in all systems, because there is such a variety of possible systems, for example, gateways, routers, terminal servers, workstations, and multi-user hosts. A separate section is devoted to the network-specific standards protocols.

B.27 Internet Protocol

Standard Name: Internet Protocol (IP)
Designation: RFC 791, 792, 919, 922, 950, 1112 for Ipv4 RFC 1883, 1884, 1885 for IPv6
Related Designations: IAB Standard 5
API: NA
Sources of Information: http://ds.internic.net/ds/rfc-index.htm

The Internet Protocol (IP) provides a connectionless service to the network layer for delivery of packets from the source to the final destination. This differs from the data link layer (Layer 2) which moves frames from one end of the wire to the other. The network layer is the lowest layer that provides end-to-end transmission. The packets at the network layer are called datagrams and contain address and routing information. Each datagram is routed independently and contains the full source and destination address.

The Internet Protocol Version 4 is defined by the IAB as IAB Standard 5 and detailed in RFCs 791,792, 919, 922, 950, and 1112. Requirements for IP Version 4 routers that apply to gateways are stated in RFC 1812. The requirements that apply to the hosts are issued as RFC 1122 and RFC 1123. The following discussions pertain to the core IP Version 4 RFCs and the related RFCs can be found in RFC 1920.

- RFC 791: Internet Protocol. This RFC defines the Internet Protocol that is designed for use in interconnected systems of packet-switched computer communication networks. The internet protocol provides for transmitting blocks of data called datagrams from sources to destinations, where sources and destination are hosts identified by fixed length addresses. The IP also provides for fragmentation and reassembly of long datagrams when transmitted through "small packet" networks. IP calls on local network protocols to carry the internet datagram to the next gateway or destination host. IP implements two basic functions: addressing and fragmentation. IP shares common rules for interpreting address fields and for fragmenting and assembling internet datagrams. The IP has been amended by RFC 919 and RFC 922 to provide IP broadcast Datagrams.

- RFC 792: Internet Control Message Protocol (ICMP) Specification. The ICMP is used when a gateway or destination host communicates with a source host, for example to report an error in datagram processing. ICMP uses the basic support of IP as if it were a higher level protocol, however, ICMP is actually an integral part of IP, and must be implemented by every IP module. ICMP messages typically report errors in the processing of datagrams. For example, this occurs when a datagram cannot reach its destination, when the gateway does not have the buffering capacity to forward a datagram, and when the gateway can direct the host to send traffic on a shorter route. Also ICMP messages are sent about errors in handling fragment zero of fragmented datagrams (fragment zero has the fragment offset equal zero).
- RFC 950: Internet Standard Subnetting Procedure. This RFC discusses the utility of "subnets" of internet networks, which are logically visible subsections of a single internet network. For administrative or technical reasons, many organizations have chosen to divide one internet network into several subnets, instead of acquiring a set of internet network numbers. This RFC specifies procedures for the use of subnets. These procedures are for hosts (e.g. workstations). The procedures used in and between subnet gateways are not fully described. Important motivation and background information for a subnetting standard is provided in RFC 940.

The Internet Protocol Version 6 (IPv6) is detailed in three RFCs (RFC 1883, 1884, and 1885). RFC 1883 (Internet Protocol Version 6 (IPv6)) is the core specification and is sometimes referred to as IP Next Generation or IPng. The focus of IPv6 over IPv4 includes providing more efficient packet handling across a global Internetwork. IPv6 is a new version of the Internet Protocol, designed as a successor to IP version 4 (RFC 791). The changes from IPv4 to IPv6 fall primarily into the following categories:

1. Expanded Addressing Capabilities. IPv6 increases the IP address size from 32 bits to 128 bits, to support more levels of addressing hierarchy, a much greater number of addressable nodes, and simpler auto-configuration of addresses. The scalability of multicast routing is improved by adding a "scope" field to multicast addresses. A new type of address called an "anycast address" is defined, used to send a packet to any one of a group of nodes.
2. Header Format Simplification. Some IPv4 header fields have been dropped or made optional, to reduce the common-case processing cost of packet handling and to limit the bandwidth cost of the IPv6 header.

3. Improved Support for Extensions and Options. Changes in the way IP header options are encoded allow for more efficient forwarding, less stringent limits on the length of options, and greater flexibility for introducing new options in the future.
4. Flow Labeling Capability. A new capability is added to enable the labeling of packets belonging to particular traffic "flows" for which the sender request special handling, such as non-default quality of service or "real-time" service.
5. Authentication and Privacy Capabilities. Extensions to support authentication, data integrity, and (optional) data confidentiality, are specified for IPv6.

This RFC specifies the basic IPv6 header and the initially-defined IPv6 extension headers and options. It also discusses packet size issues, the semantics of flow labels and priority, and the effects of IPv6 on upper-layer protocols. The format and semantics of IPv6 addresses are specified separately in RFC 1884. The IPv6 version of ICMP, which all IPv6 implementations are required to include, is specified in RFC 1885.

B.28 Integrated Services Digital Network

Standard Name: Integrated Services Digital Network (ISDN)
Designation: FIPS 182
Related Designations: None
API: NA
Sources of Information:
- http://ISDN.NOSL.NIST.gov/misc/NIUF.html
- http://www.NIUF.NIST.gov

FIPS 182 Integrated Services Digital Network (ISDN) defines the generic protocols necessary to establish transparent ISDN connections among government networks and between government and conformant common carrier networks. FIPS 182 supports a range of integrated services including voice, data, image, and video services. It is comprised of the protocols, Implementation Agreements and conformance tests regarding the D-channel procedures for the underlying (Layers 1, 2, and 3) ISDN protocols, as well as a limited set of other protocols, such as ISDN Bearer Services, X.25 Packet Services, and Terminal Adaptation. This standard is consistent with FIPS 146-2 POSIT, which provides protocols for computer to computer data communications using ISDN as a lower layer network technology.

B.29 Transmission Control Protocol

Standard Name:	Transmission Control Protocol (TCP)
Designation:	RFC 793
Related Designations:	IAB Standard 7
API:	NA
Sources of Information:	http://ds.internic.net/ds/rfc-index.htm

The Transmission Control Protocol (TCP) is intended for use as highly reliable host-to-host protocol between hosts in packet-switched computer communication networks, and in interconnected systems of such networks. TCP was originally formulated for use on the ARPANET. It was designed to tolerate an unreliable subnetwork. The TCP transport service works by accepting varying sizes of messages from users, and breaking them up into pieces that do not exceed 64 KB. The pieces are then sent through the network as separate datagrams. The protocol is responsible for the tracking and retransmission of undelivered datagrams. It is also responsible for reassembling the datagrams after they arrive at the destination.

TCP is a connection-oriented, end-to-end reliable protocol designed to fit into a layered hierarchy of protocols which support multi-network applications. The TCP provides for reliable inter-process communication between pairs of processes in host computers attached to distinct but interconnected computer communication networks. Very few assumptions are made as to the reliability of the communication protocols below the TCP layer. TCP assumes it can obtain a simple potentially unreliable datagram service from the lower level protocols. In principle, the TCP should be able to operate above a wide spectrum of communication systems ranging from hard-wired connections to packet-switched or circuit-switched networks.

B.30 Videoconferencing

Standard Name:	Narrow-Band Visual Telephone Systems and Terminal Equipment
Designation:	ITU-T H.320
Related Designations:	None
API:	NA
Sources of Information:	Contact ITU for the specification · http://www2.echo.lu/oii/en/confer.html#MMC • http://www.itu.ch/itudoc/itu-t/rec/h/h320_23397.html

The H.320 family of standards forms the basis for communication between desktop conferencing systems between different suppliers. The H.320 family of standards have been developed specifically for videoconferencing and imposes no particular conditions on the applications for handling facilities such as e-mail, file transfer, directory, etc., which are typically bundled into desktop products. At the heart of the H.320 standards is H.261 Video Codec for Audiovisual Services at px64 kbps, where p'1,2...30. Like JPEG and MPEG, H.261 incorporates the Discrete Cosine Transform (DCT) algorithms, thus making it possible for the development of truly multimedia processors and multifunctional codices. The ITU-T has adopted two formats¾Common Intermediate Format (CIF) at 352x288 and Quarter-CIF Format (QCIF) at 176x144¾as the video formats for videoconferencing/video telephony. All codecs must be able to operate with QCIF, whilst CIF is optional. CIF is more suitable for higher bandwidth conferencing facilities such as videoconferencing (where p is greater than or equal or 6), whereas QCIF is more suitable for lower bandwidth conferencing facilities such as video telephony, where the value of p is generally lower.

A new suite of standards, known as the H.324 suite, is currently being prepared for low bit rate multimedia communications. These standards will be based on the H.263 codec rather than H.261. H.263 considerably broadens H.261 capabilities, extending from low-resolution, credit card-size, images to high-quality, exceptional detail images, for applications such as telemedicine. Backward compatibility to H.320 QCIF (Quarter Common Intermediate Format) and, optionally, to CIF and SQCIF (Sub-QCIF) is provided.

The H.320 standards have been widely implemented into today's desktop conferencing systems and can potentially revolutionize the PC desktop market, especially with the expected arrival of multimedia chipsets. However, some implementations do not support the full set of H.320 standards. Options, enhancements and extensions to the standards leave considerable scope for suppliers to gain competitive edge and could pose problems for interoperability between equipment from different suppliers. A number of suppliers have developed specific codecs to support the H.320 standards.

At the time of this writing the following standards comprised the H.320 family: H.200, H.221 Rev 2, H.230 Rev 2, H.231, H.233, H.242 Rev 2, H.243, H.261 Rev 2, H.263, H.320 Rev 2, H.321, H.322, H.323, H.324, H.331, G.711, G.722, G.722 Annex a, G.728, G.172, F.730, F.740. The reader should review each of these standards for relevance to their application.

B.31 LDAP

Standard Name: Lightweight Directory Access Protocol
Related Designations: DAP (Directory Access Protocol), X-500
Sources of Information: Internet Directory Consortium http://www.opengroup.org/idc/
- Innsoft International LDAP Web site: http://www.critical-angle.com/ldapworld/index.html

A client-server protocol for accessing a directory service. It was initially used as a front-end to X.500, but can also be used with stand-alone and other kinds of directory servers. LDAP is under IETF change control and so can evolve to meet internet requirements. It is expected that

LDAP will provide a common method for searching e-mail addresses on the internet, eventually leading to a global white pages.

LDAP is a simplified version of the DAP protocol, which is used to gain access to X.500 directories. It is easier to code the query in LDAP than in DAP, but LDAP is less comprehensive. For example, DAP can initiate searches on other servers if an address is not found, while LDAP cannot in its initial specification.

B.32 Internet Calendar Access Protocol

Designation: ICAP
Related Designations: iCalendar
Sources of Information: Internet Mail Consortium http://www.imc.org/pdi/

Internet Calendar Access Protocol (ICAP) allows a client to access, manipulate and store Calendar information on a server. ICAP employs the iCalendar format for interchange of calendaring. ICAP includes operations for creating Calendar stores on a server, opening them and retrieving information about them. When a Calendar Store has been opened, it can be bounded by start and end times so that the client can act on a smaller subset of Calendar information for more efficient operation. ICAP allows users to store new Calendar Objects into their own Calendar store and Calendar stores owned by other users with a single operation.

ICAP supports searching iCalendar objects within a Calendar Store. Searches can be based on any iCalendar property and filtered by iCalendar Component type.

Beginning in December, 1996, the Internet Mail Consortium took on responsibility for the development and promotion of iCalendar technologies.

The IESG has approved the specification for iCalendar as proposed standards. The RFCs are

- RFC 2445, Internet Calendaring and Scheduling Core Object Specification (iCalendar)
- RFC 2446, iCalendar Transport-Independent Interoperability Protocol (iTIP): Scheduling Events, BusyTime, To-dos and Journal Entries
- RFC 2447, iCalendar Message-based Interoperability Protocol (iMIP)

B.33 vCalendar

Related Designations: iCalendar
Sources of Information: Internet Mail Consortium http://www.imc.org/pdi/

vCalendar is a Personal Data Interchange (PDI) specification developed by Versit and endorsed by several manufacturers, including Microsoft Outlook 98, Netscape Communicator 4, Professional Edition and Lotus Organizer 97 GS.

vCalendar defines a format for exchanging calendaring and scheduling information. It captures information about event and to-do items that are used by applications such as PLMs and group schedulers. Programs that use vCalendar can exchange data about events so that you can schedule meetings with anyone who has a vCalendar-aware program.

B.34 Metadata Interchange Specification (MDIS)

Standard Name:	Metadata Interchange Specification
Designation:	MDIS 1.0
Related Designations:	OIM, Metadata Coalition
API:	MDIS 1.0; also file format definitions and profile definitions.
Sources of Information:	http://www.MDCinfo.com; http://www.metadata.org

The Metadata Interchange Specification Initiative brings industry vendors and users together to address a variety of problems and issues regarding the exchange, sharing, and management of metadata. This is a voluntary coalition of interested parties with a common focus and shared goals, not a traditional standards body or regulatory group.

To enable full-scale enterprise data management, different IT tools must be able to freely and easily access, update, and share metadata. The only viable mechanism to enable disparate tools from different vendors to exchange metadata is a common metadata interchange specification with guidelines to which the different vendors' tools can comply.

In choosing the interchange-compliant tools, purchasers can be assured of the accurate and efficient exchange of metadata essential to meeting their users' business information needs. This will allow IS managers to build on investments in data management tools and infrastructure with each additional product purchase.

The goal of the Metadata Interchange Specification (MDIS) is to provide a specification that can be implemented within a two- to four-person effort by the average vendor.

This is not intended as a typical standards specification effort, where the goal is to create a standard definition of all the possible information pertinent to the domain and the format for representing it. The assumption here is that for some period of time, at least, the contents of what is considered metadata will be in flux. The most important goal of the MDIS is to define an extensible mechanism that will allow vendors to exchange common metadata as well as carry along "proprietary" metadata.
The founding members agreed upon initial short-term goals, including

- Creating a vendor-independent, industry-defined and -maintained standard access mechanism and standard application programming interface (API) for metadata;

- Enabling users to control and manage the access and manipulation of metadata in their unique environments through the use of interchange specification-compliant tools;
- Allowing users to build tool configurations that meet their needs and to incrementally adjust those configurations as necessary to add or subtract tools without impact on the interchange specification environment;
- Enabling individual tools to satisfy their specific metadata access requirements freely and easily within the context of an interchange model;
- Defining a clean, simple interchange implementation infrastructure that will facilitate compliance and speed adoption by minimizing the amount of modification required to existing tools to achieve and maintain MDIS compliance; and
- Creating a process and procedure for establishing and maintaining the MDIS and for extending and updating it over time as required by evolving industry and user needs.

B.35 Open Information Model (OIM)

Standard Name:	Open Information Model
Designation:	OIM
Related Designations:	MDIS, Metadata Coalition
Sources of Information:	http://www.MDCinfo.com

Part of the Microsoft Repository, the OIM is a metadata model and specification, a product that provides metadata management services. The metadata model was developed by Microsoft, together with more than 20 industry-leading companies, and has been reviewed by more than 300 companies as part of Microsoft's Open Process. Microsoft, a new member of the Metadata Coalition has announced the transfer control of the Open Information Model to that consortium. The Metadata coalition will maintain and evolve the Open Information Model as a technology-independent and vendor-neutral metadata standard.

OIM is a set of metadata specifications to facilitate sharing and reuse in the application development and data warehousing domains. OIM 1.0 consists of over 200 types organized in easy-to-use and easy-to-extend subject areas.

The representation of OIM instances based on XML provides a powerful and easy-to-implement mechanism to exchange metadata between multiple heterogeneous repositories offered by different vendors. This is the first time that enterprise customers are able to interchange and integrate application development and data warehousing metadata using a standard encoding format and a broadly accepted industry standard information model.

B.36 The Metadata Coalition

Standard Name: Metadata Coalition
Designation: MDC
Related Designations: MDIS
Sources of Information: http://www.MDCinfo.com

The Metadata Coalition (MDC), founded in 1995, is a not-for-profit consortium of vendors and end-users whose goal is to provide a tactical solution for metadata exchange (MDIS). The consortium includes dozens of vendors of enterprise tools, metadata management and data warehousing products. Participation in the MDC is encouraged and open to all vendors and end users. The MDC Council comprises Commercial Financial Services, Inc., ETI, IBM, NCR, PLATINUM technology, inc., and Sybase. The Technical Subcommittee currently comprises PriceWaterhouse-Coopers, NCR, IBM, Sybase, Cognos, SAS, CFS, Platinum Technology, Inc., One Meaning, Mastersoft International (MSI) and Prudential.

Microsoft become a member of the Metadata Coalition in December 1998 with the goal of helping solidify the Coalition's role in evolving a technology and vendor-independent standard for metadata interchange. Microsoft will provide an implementation-independent version of its Open Information Model (OIM) to the Coalition and work as a member of the technical subcommittee to devise a plan for merging the OIM with the Coalition's Metadata Interchange Specification (MDIS). This makes the Coalition the long-term owner of a unified industry standard for metadata that already is in widespread productive use and guarantees that it evolves in a vendor-neutral and upward-compatible fashion.

"Companies need standards to help them share information models and metadata in a heterogeneous, distributed environment," stated Bob Craig of the Hurwitz Group. "This announcement is an important milestone in establishing a standards-based foundation for metadata sharing between enterprise applications and databases. Hurwitz believes the Metadata Coalition is performing an important service which will ultimately benefit the end user."

The results of this collaboration will provide the enterprise market with

- A technology-independent and vendor-neutral information model describing the structure and semantics of metadata,
- An implementation-independent XML-based interchange format for metadata, and
- A platform for vendors and end-users to collaborate on the design of the above.

B.37 XML Interchange Format (XIF)

Standard Name:	XML Interchange Format
Designation:	XIF
Related Designations:	XML, OIM
Sources of Information:	World Wide Web Consortium (W3C) http://www.w3.org/

The XML Interchange Format for Cross-Repository Metadata Exchange is designed to support the technology-independence of the Open Information Model. The XML Interchange Format allows metadata to be moved between any two repository products. Key features of XIF include

- XIF is technology-independent—It is based on XML, a W3C format for the representation of information as structured documents. The format is technology independent, published, and easy to use.
- XIF is vendor-independent—Most vendors of repository products have announced support for XIF. The vendors planning to support XIF by early 1999 are Microsoft, NCR, PLATINUM technology, Siemens Nixdorf, Softlab, Sybase, UNISYS, and Viasoft.
- XIF works with the Open Information Model—As an open, industry-standard model accommodating metadata of software development and data warehousing tools, the Open Information Model (OIM) provides a content-rich, yet vendor-neutral specification of metadata. Vendors supporting XIF will be able to import and export metadata, such as analysis and design models, component descriptions, and data warehousing transformations.
- XIF Import/Export Utility Available—The Microsoft Repository SDK Version 2.1 includes the Microsoft Repository XML Import/Export Utility and is available for download.
- XIF provides third-party vendors with an easy way to populate repository databases with data. For example, a third-party tool can insert OIM-compliant instances into Microsoft Repository by creating an XML file, then using the Microsoft XIF Import/Export Utility.

B.38 H.323

Internet Engineering Task Force (IETF) specifications Industry standard for audio and video communications over the internet.

An ITU standard for videoconferencing over packet-switched networks such as local area networks (LANs) and the internet. It allows any combination of voice, video and data to be transported. H.323 specifies several video codecs, including H.261 and H.263, and audio codecs, including G.711 and G.723.1. Gateways, gatekeepers and multipoint control units (MCUs) are also covered. H.323 is widely supported for internet telephony.

B.39 T.120

An ITU standard for real-time data conferencing (sharing data among multiple users). It defines interfaces for whiteboards, application viewing and application sharing. The ITU standard for videoconferencing is H.320.

T.120 is an umbrella term for a series of specifications that define all aspects of data conferencing. The complete standard is made up of the following components.

B.40 PICS

The W3C Digital Signature Working Group ("DSig") proposed standard format for making digitally-signed, machine-readable assertions about a particular information resource. PICS 1.1 labels are an example of such machine-readable assertions. This document describes a method of adding extensions to PICS 1.1 labels for purposes of signing them. More generally, it is the goal of the DSig project to provide a mechanism to make the statement: signer believes statement about information resource. In DSig 1.0 statement is any statement that can be expressed with PICS 1.1."

B.41 Extensible Markup Language (XML)

Standard Name:	Extensible Markup Language
Designation:	XML
API:	Yes
Sources:	W3C: http://www.w3.org/RDF/

XML defines a universal standard for electronically exchanging data. XML specifies a rigorous, text-based way to represent the structure inherent in data so that it can be authored and interpreted unambiguously. Its simple, tag-based approach leverages developers' familiarity of HTML but provides a flexible, extensible mechanism that can handle the gamut of "digital assets" from highly structured database records to unstructured documents.

XML is an Internet Standard way of tagging data. As a Web-centric subset of the well-respected SGML standard, XML is based on a proven technology with a good track record. The Worldwide Web Consortium (W3C) recommended the XML 1.0 standard in February, 1998, and it is being widely and rapidly adopted as a standard for document and data exchange in a variety of markets. XML is gaining wide industry support as well from vendors like Oracle, IBM, Sun, Microsoft, Netscape, SAP and others, as a platform- and application-neutral format for exchanging information.

Companies such as Oracle have begun to participate actively in steering the standards to directly influence the evolution of XML. At present, Oracle is working together with IBM, Microsoft, and others in the W3C XML Working Group to define an Internet Standard for "XML Schemas". This effort will create a standard to describe the structure and data types of a document's elements to enable more automated and seamless integration of XML with databases and programming languages in the future.

B.42 Document Object Model (DOM)

Standard Name: Document Object Model Level 1
Designation: DOM
Related Designations: HTML, XML

The Document Object Model Level 1, is a platform- and language-neutral interface that allows programs and scripts to dynamically access and update the content, structure and style of documents. It was endorsed by the W3C as a W3C Recommendation in October 1998. The Document Object Model provides a standard set of objects for representing HTML and XML documents, a standard model of how these objects can be combined, and a standard interface for accessing and manipulating them. Vendors can support the DOM as an interface to their proprietary data structures and APIs, and content authors can write to the standard DOM interfaces rather than product-specific APIs, thus increasing interoperability on the Web.

The goal of the DOM specification is to define a programmatic interface for XML and HTML. The DOM Level 1 specification is separated into two parts: Core and HTML. The Core DOM Level 1 section provides a low-level set of fundamental interfaces that can represent any structured document, as well as defining extended interfaces for representing an XML document. These extended XML interfaces need not be implemented by a DOM implementation that only provides access to HTML documents; all of the fundamental interfaces in the Core section must be implemented. A compliant DOM implementation that implements the extended XML interfaces is required to also implement the fundamental Core interfaces, but not the HTML interfaces. The HTML Level 1 section provides additional, higher-level interfaces that are used with the fundamental interfaces defined in the Core Level 1 section to provide a more convenient view of an HTML document. A compliant implementation of the HTML DOM implements all of the fundamental Core interfaces as well as the HTML interfaces.

B.43 SAX 1.0: The Simple API for XML

Standard Name:	The Simple API for XML
Designation:	SAX 1.0
Related Designations:	XML
Sources of Information:	http://www.megginson.com/index.html

SAX is an emerging de facto public-domain standard interface for event-based XML parsing. Developed collaboratively by the members of the XML-DEV mailing list. SAX 1.0 was released on Monday 11 May 1998, and is free for both commercial and non-commercial use. SAX implementations are currently available in Java and Python, with more to come. SAX 1.0 support in both parsers and applications is growing fast.

B.44 USMARC

Standard Name:	United States Machine Readable Cataloging
Designation:	USMARC
Related Designations:	ANSI Z39.2, ISO 2709
API:	Utilities exist for converting between text and MARC—MARCMakr, MARCBreakr
Sources of Information:	http://lcWeb.loc.gov/marc/index.html

The Network Development and MARC Standards Office is a center for library and information network standards and planning in the Library of Congress. Established in 1976 to provide focus for networking activities in the Library of Congress, the office was expanded in 1984 to include MARC standards responsibilities. Staff are involved in many facets of network development including

- Standards, which are basic to efficient, long-term interchange with other systems such as those for Machine-Readable Cataloging (MARC) and Z39.50 information retrieval protocols;
- Planning, which involves working out detailed models and specifications with other institutions and with internal Library of Congress units; and
- Coordinating and testing implementation that takes the standards development and planning to fulfillment through the completion of operational networking systems.

The USMARC formats are standards for the representation and communication of bibliographic and related information in machine-readable form. A USMARC record involves three elements: the record structure, the content designation, and the data content of the record.

The structure of USMARC records is an implementation of national and international standards, e.g., Information Interchange Format (ANSI Z39.2) and Format for Information Exchange (ISO 2709). A USMARC format is a set of codes and content designators defined for encoding machine-readable records. Formats are defined for five types of data: bibliographic, holdings, authority, classification, and community information.

USMARC Format for Bibliographic Data contains format specifications for encoding data elements needed to describe, retrieve, and control various forms of bibliographic material. The USMARC Format for Bibliographic Data is an integrated format defined for the identification and description of different forms of bibliographic material. USMARC specifications are defined for books, serials, computer files, maps, music, visual materials, and mixed material. With the full integration of the previously discrete bibliographic formats, consistent definition and usage are maintained for different forms of material. The USMARC formats are maintained by the Library of Congress in consultation with various user communities.

The USMARC formats are communication formats, primarily designed to provide specifications for the exchange of bibliographic and related information between systems. They are widely used in a variety of exchange and processing environments. As communication formats, they do not mandate internal storage or display formats to be used by individual systems.

B.45 Dublin Core

Standard Name:	The Dublin Core: A Simple Content Description Model for Electronic Resources
Designation:	Dublin Core
Related Designations:	NA
API:	NA
Sources of Information:	http://purl.oclc.org/dc/

The Dublin Core is a metadata element set intended to facilitate discovery of electronic resources. Originally conceived for author-generated description of Web resources, it has attracted the attention of formal resource description communities such as museums, libraries, government agencies, and commercial organizations.

The Dublin Core is intended to be usable by non-catalogers as well as resource description specialists. Most of the elements have a commonly understood semantics of roughly the complexity of a library catalog card. Disparate description models interfere with the ability to search across discipline boundaries. Promoting a commonly understood set of descriptors that helps to unify other data content standards increases the possibility of semantic interoperability across disciplines.

Recognition of the international scope of resource discovery on the Web is critical to the development of effective discovery infrastructure. The Dublin Core benefits from active participation and promotion in 20 countries in North America, Europe, Australia, and Asia.

The Dublin Core provides an economical alternative to more elaborate description models such as the full MARC cataloging of the library world. Additionally, it includes sufficient flexibility and extensibility to encode the structure and more elaborate semantics inherent in richer description standards

The diversity of metadata needs on the Web requires an infrastructure that supports the coexistence of complementary, independently maintained metadata packages. The World Wide Web Consortium (W3C) has begun implementing an architecture for metadata for the Web. The Resource Description Framework, or RDF, is designed to support the many different metadata needs of vendors and information providers.

Representatives of the Dublin Core effort are actively involved in the development of this architecture, bringing the digital library perspective to bear on this important component of the Web infrastructure. Dublin Core Metadata sponsors include

- CNI (Coalition for Networked Information)
- DSTC (Distributed Systems Technology Centre)
- LC (The Library of Congress)
- NLA (National Library of Australia)
- NLF (National Library of Finland)
- NSCA (National Computational Science Alliance)
- NSF (National Science Foundation)
- OCLC (Online Computer Library Center)
- UKOLN (UK Office for Library and Information Networking)

B.46 STEP

Standard Name: Standard for the Exchange of Product Data
Designation: STEP
Related Designations: ISO standard (ISO 10303)
API: SDAI
Sources of Information: PDES Inc. http://pdesinc.scra.org/
STEPtools Inc. http://www.steptools.com/library/standard/

STEP is designed to cover all information through a product's life cycle. STEP includes standard formats (APs), a standard language (Express) and standard APIs (SDAI). STEP is a set of ISO standards which provide for the exchange of engineering product data. These standards can be grouped into infrastructure components and industry specific information models:

- The EXPRESS information modeling language (Part 11)
- An EXPRESS-driven data exchange file specification (Part 21)
- An EXPRESS-driven application programming interface (SDAI) with bindings to the C, C++, and IDL languages (Parts 22-26)
- A conformance testing framework (Part 31)
- A library of general purpose information models for things like geometry, topology, product identification, dates, times, etc. (The 40-series parts)
- Industry-specific application protocols that are built from the library of general models (the 200-series parts):
- Explicit Drafting (Part 201)
- Associative Drafting (Part 202)
- Configuration Controlled 3D Assemblies (Part 203)
- Electronic Chip and Printed Circuit Board Design (Part 210)
- Structural Analysis (Part 209)
- Cabling (Part 212)
- Facilitates (Part 227)

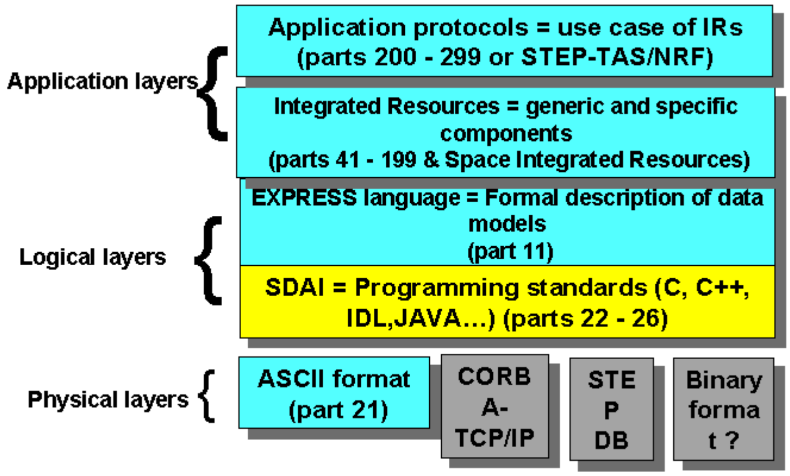


Figure B-5. STEP: A Complete Product Information Architecture. An ISO Standard (10303) Distributed In Parts.

B.47 RDF

Standard Name: Resource Description Framework
Designation: RDF
Related Designations: XML
Sources of Information: World Wide Web Consortium (W3C)
http://www.w3.org/RDF/

The Resource Description Framework (RDF) is a specification currently under development within the W3C Metadata activity. RDF is designed to provide an infrastructure to support metadata across many Web-based activities. RDF is the result of a number of metadata communities bringing together their needs to provide a robust and flexible architecture for supporting metadata on the internet and WWW. Example applications include sitemaps, content ratings, stream channel definitions, search engine data collection (Web crawling), digital library collections, and distributed authoring.

RDF will allow different application communities to define the metadata property set that best serves the needs of each community. RDF will provide a uniform and interoperable means to exchange the metadata between programs and across the Web. Furthermore, RDF will provide a means for publishing both a human-readable and a machine-understandable definition of the property set itself.

RDF will use XML as the transfer syntax in order to leverage other tools and code bases being built around XML. The following documents have been released for public comment

- A revised public draft of the RDF Model and Syntax Specification was released on October 11, 1998—Proposed Recommendation due during December 1998.
- A revised public draft of the RDF Schema Specification was released on November 3, 1998—Proposed Recommendation during January 1999.

The W3C RDF Working Group has key industry players including DVL, Grif, IBM, KnowledgeCite, LANL, Microsoft, Netscape, Nokia, OCLC, Reuters, SoftQuad and University of Michigan.

The RDF Working Group is one of the earliest phases of a major effort by the Consortium to build a vendor-neutral and operating system- independent system of metadata. The collaborative design effort on RDF originated as an extension on the PICS content description technology, and draws upon the XML design as well as recent W3C Submissions by Microsoft [XML Web Collections] and Netscape [XML/MCF]. In addition, documents such as Microsoft's XML-Data and Site Map proposals, and the Dublin Core/Warwick Framework have also influenced the RDF design.

RDF will allow different application communities to define the metadata property set that best serves the needs of each community. RDF metadata can be used in a variety of application areas such as

- In resource discovery to provide better search engine capabilities;
- In cataloging for describing the content and content relationships available at a particular Web site, page, or digital library;
- By intelligent software agents to facilitate knowledge sharing and exchange;
- In content rating for child protection and privacy protection;
- In describing collections of pages that represent a single logical "document";
- For describing intellectual property rights of Web pages.

With digital signatures, RDF will be key to building the "Web of Trust" for electronic commerce, collaboration, and other applications.

C Governing Policies and Procedures

JPL, under contract to NASA is a federally funded scientific and research organization whose fundamental purpose is to create knowledge and to share that knowledge freely with the scientific community and the public. JPL also has the responsibility to protect its IT resources in order to minimize, and where possible, eliminate threats to loss of knowledge-based assets. At the same time, the US government and JPL's agreements with industry partners require JPL to restrict some information from some people (for example, ITAR information from foreign nationals). The sections below discuss many of the legal or contractual vehicles that levy requirements on JPL's implementation of knowledge management.

C.1 Prime Contract

JPL's Prime Contract (NAS7-1407) provides a number of contractual obligations and goals that apply to knowledge management. One of the objectives of the knowledge management effort at JPL is to maximize emphasis on knowledge management activities that satisfy both JPL internal needs and simultaneously help meet contractual commitments. Examples of other applicable policies and guidelines follow.

C.1.1 JPL Policy—Releasing Information Outside of JPL

This policy requires approval by the cognizant manager and by Document Review Services for information released outside of JPL, including release to contracting partners and affiliates. The only information items exempted from this are specific classes of information released by the procurement division, Media Relations Office, and Educational Affairs Office. The purpose of this policy is to ensure compliance with International Traffic and Arms Regulations (ITAR), Export Administration Regulations (EAR), and intellectual property rights for Caltech and JPL's partners.

The knowledge management process owner will be responsible for clarifying this policy and data release requirements for project managers and staff in conjunction with the existing JPL network architecture and user identification techniques available.

C.1.2 Institutional Configuration Management (Category A and B) Policy

JPL currently has an institutional configuration management policy that is undergoing review and revision as part of the ISO working groups' effort to streamline and clarify existing policies. The new policy should be reviewed and worked into each of the knowledge management services that require consistency with this policy.

C.1.3 ISO 9001

JPL is required to become ISO 9001 certified during 1999 as part of an overall NASA Agency initiative. Four of the ISO 9001 elements have a close relationship with JPL's overall knowledge management architecture, requirements, and solutions. Working groups established to prepare JPL for upcoming ISO audits include representation from the KM Study Team. Most of these policies are under review and may change to reflect new requirements. All of the resulting policies should be reviewed closely by the knowledge management implementation team and factored in to detailed design and implementation decisions during knowledge management service development.

C.1.4 JPL Policy—Management of Inactive and Archival Institutional Records

This draft policy defines active, inactive, and archival records. It includes policies for storage and maintenance of these records for long-term access and preservation. The Prime Contract also contains requirements for the archive of records.

C.1.5 NASA Standards

The NASA Technical Standards Program [68] is designed to support the needs of all NASA enterprises and JPL as part of the cross-cutting processes identified in the NASA Strategic Management Plan. As part of JPL's prime contract, several of these standards are identified as recommended goals for JPL to ensure interoperability across the Agency and with JPL's affiliates. The implementation team should take an active role in the evolution and adoption of these standards at JPL to further both NASA's and JPL's knowledge sharing objectives.

D Glossary

Term	Description
Ad hoc query	A query that cannot be determined before the query is made (ATG).
Aggregation	The process whereby a *data value* is derived from the collection of different data occurrences of the same subject data (ATG).
Application program interface (API)	An interface of a system at which services are provided to application software (the term application is used here to characterize the function using the service) (ISO/IEC TR 10000-3). When a software system features an API, it provides a means by which programs written outside of the system can interface with the system to perform additional functions. For example, a data mining software system may have an API that permits user-written programs to perform such tasks as extract data, perform additional statistical analysis, create specialized charts, generate a model, or make a prediction from a model.
Archive	A storage system designed for long-term access and preservation of material or information.
Attribute	A characteristic or element of an entity to which a value can be assigned for the purpose of capturing/providing vital information about that entity, e.g., title, author, date, reference designator; also referred to as data value type or element. Are classified as either general class (e.g., date) or subclass (e.g., publication date).
Building block	A modular component that can be used to implement a knowledge resource. Building blocks exist at all levels of detail and sophistication and may be used to create larger building blocks. Examples of building blocks are an API, a template, a software utility, and a Project library.
Capture	The process by which data and metadata is input to an online *information management system*.
Catalog	A component of a data dictionary containing a directory of its database entities as well as the attributes of each entity (ATG).
Categorical data	Data that fits into a small number of discrete categories (as opposed to continuous) either as non-ordered (nominal) such as gender or city, or as ordered (ordinal) such as high, medium, or low temperatures.
Change control	A process that ensures all changes are properly identified, reviewed, approved, implemented, tested, and documented (DOE-1073-93 11/93).
Class	A named category of items that share common characteristics and behavior (ISO 10303).
Cleansing	The process of removing errors and resolving inconsistencies in source data before loading the data into a target environment (ATG).
Communications services interface	An interface of a system at which interactions take place with entities external to the system, such as external data transport facilities and functions in other systems (ISO/IEC TR 10000-3).
Compound document	A set of files that consists of a master (parent or top) file and one or more sub-files (slave, child, or subordinate) that are related to the master file (JPL D-14206).

Term	Description
Container	An organized, structured collection of information that meets a specific need of a set of users; containers have the following attributes associated with them: purpose, owner, users, content, context, policy, (7) Operations Concept, (8) Meaningful Measures, (9) Structure, (10) Access Method
Data	A representation of information in a formed manner suitable for communication, interpretation, or processing by a human being or computer (ISO 10303).
Database	A collection of data that are logically related (ATG).
Database management system (DBMS)	A software system for creating, maintaining, and protecting databases (ATG).
Data dictionary	Contains attribute definitions (general *class*), subclass definitions, definition guidelines for each attribute (class and subclass), and entity definitions; includes a *catalog*.
Data exchange	The storing, accessing, transferring, and archiving of data (ISO 10303).
Data file	An organized collection of digital data representing a unit or units of information, such as a document, drawing, spreadsheet, form, or any part thereof, that is created by a specific software package; also referred to as a *file.*
Data mart	A type of data warehouse designed to meet the needs of a specific group of users, such as a department or part of an organization; usually focuses on a single subject area (ATG).
Data mining	An information extraction activity whose goal is to discover hidden facts, relationships, patterns, and associations contained in databases. Using a combination of machine learning, statistical analysis, modeling techniques, and database technology, data mining finds patterns and subtle relationships in data and infers rules that allow the prediction of future results. Typical applications include market segmentation, customer profiling, fraud detection, evaluation of retail promotions, and credit risk analysis.
Data model	A logical map that represents the inherent properties of the data independent of software, hardware, or machine performance considerations (ATG); see also *model* and *enterprise data model.*
Data producer	The individual primarily responsible for creating the information; content provider; author; creator.
Data set	Any defined collection of data; usually refers to the data maintained within a specified information system, including all files and metadata.
Data specification language	A set of rules for defining data and their relationships suitable for communication, interpretation, or processing by computers (ISO 10303).
Data value	A value assigned to an *attribute* or data value type; e.g., 1998-10-04 is one of the values assigned to the attribute named (labeled) “date.” An attribute usually has any number of values assigned to it. In a relational database, the values assigned to any one attribute comprise a *field.*
Data value type	A characteristic or element of an entity to which a value can be assigned for the purpose of capturing/providing vital information about that entity, e.g., title, author, date, reference designator; also referred to as data value type or element. Are classified as either general class (e.g., date) or subclass (e.g., publication date).

Term	Description
Data warehouse	A center of the architecture for information systems in the 1990s; supports informational processing by providing a solid platform of integrated, historical data from which to do analysis; provides the facility for integration in a world of non-integrated application systems; organizes and stores the data needed for informational, analytical processing over a long historical time perspective; collects data in support of management's decision-making process; is subject-oriented, integrated, time-variant, nonvolatile.
Decision-support system (DSS)	A system that allows decision-makers to access data relevant to the decisions they are required to make (ATG).
Dissemination	The process by which information moves from the provider to the user; involves all the mechanisms that move the information in an efficient manner while complying with all applicable regulations (JPL D-14206).
Drill down	A method of exploring detailed data used in creating a summary level of data; drill down levels depend on the granularity of the data (ATG).
Element	A characteristic or element of an entity to which a value can be assigned for the purpose of capturing/providing vital information about that entity, e.g., title, author, date, reference designator; also referred to as data value type or element. Are classified as either general class (e.g., date) or subclass (e.g., publication date).
End-user	A consumer of information; also referred to as knowledge worker or user.
Enhanced data	Data that has been cleansed, transformed, and/or summarized (ATG).
Enterprise data model	A blueprint for all the data used by all the organizations in the enterprise. The model resolves all the potential inconsistencies and interpretations of the data and presents a commonly understood view and definition of the enterprise data.
Entity	One of the following: (1) a unit or units of information that is represented by a *file*, (2) an *object*, (3) a system, process, structure, or design that is represented by a *model*.
Exchange data set	A coherent and valid set of instances of entities. The data set is instantiated in the form of a physical file, a working form, or a shared database (ISO 10303).
Exchange structure	A computer-interpretable format used for storing, accessing, transferring, and archiving data (ISO 10303).
Extensible	An architecture or a system using an architecture that provides for scaling up or down in response to changing requirements and environment.
Extraction	The process of copying a subset of data from a source to a target environment (ATG).
Field	Contains the values assigned to any one attribute for all entities in a database table.
File	An organized collection of digital data representing a unit or units of information, such as a document, drawing, spreadsheet, form, or any part thereof, that is created by a specific software package; also referred to as a *data file.*
File extension	A three- or four-character suffix appended to a filename that indicates the file type, e.g., afile.doc or index.html (JPL D-14206); see also *MIME type*.
Granularity	Refers to the level of detail contained in the data.
Human/computer interface	An interface of a system at which physical interactions take place between a human being and the system (ISO/IEC TR 10000-3); a.k.a. user interface.

Term	Description
Indexing	Creation of an arranged list of specified data (JPL D-12950).
Information	Facts, concepts, or instructions (ISO 10303).
Information management system	Characterized by methods and mechanisms for capturing, managing, and delivering electronic information; examples include the various types of databases and searchable Web sites (JPL D-14206); a.k.a. information system.
Information model	A formal model of a bounded set of facts, concepts, or instructions to meet a specified requirement (ISO 10303).
Information services interface	An interface of a system at which interactions take place with external persistent storage (e.g., removable disk storage) (ISO/IEC TR 10000-3).
Instance	A single, specified example of a class or object.
Knowledge assets	The knowledge regarding markets, products, technologies, and organizations that a business owns or needs to own and which enable its business processes to generate profits (AIAI).
Knowledge base	1. A body of knowledge organized to meet a specific set of needs. Examples: a person, a set of books, a database. 2. The set of all knowledge resources within a particular domain. The *JPL Knowledge Base* is the complete set of all JPL knowledge resources. There may be subject or discipline-specific knowledge bases as well that are subset of the JPL Knowledge Base, targeted to a specific community of users.
Knowledge management	The identification and analysis of available and required knowledge, and the subsequent planning and control of actions to develop knowledge assets so as to fulfill organizational objectives (AIAI).
Knowledge resource	A structured set of information, systems, services, and contextual information that exists for an explicit purpose to meet the needs of specified users. A knowledge resource can be of personal or subject-specific value. Information in the resource has typically been culled or refined to provide added-value, and human resources or stewards exist to provide additional context as needed for the knowledge resource. (A customized view or query into a more general knowledge repository could be considered a type of knowledge resource.)
Knowledge repository	An organized collection of information stored for planned or potential future use. Examples include a database, a electronic data warehouse, a physical warehouse of records, a file cabinet, photos, vellum files, microfiche. A knowledge repository is often archival in nature with a focus on collection and long-term storage.
Knowledge worker	A consumer of information; also referred to as user or end-user.
Label	The formal, single-word name given to an attribute (Dublin).
Legacy data	Data (either digital or otherwise) that resides or used to reside in an outdated repository and has been loaded or is planned to be loaded into a current system; legacy data usually needs to be enhanced prior to loading.
Load	The process of populating a database or data warehouse. Loading is usually done using vendor-supplied utilities or user-written programs.
Meta catalog	A metadata dictionary that encompasses all the data information across the enterprise; includes physical location of data, relationships and hierarchies, and business rules (Telos); a.k.a. meta repository.

Term	Description
Metadata	Data about data; simple, structured data used to describe complex, unstructured data. Includes both definitions and relationships (JPL D-12950).
Middleware	A communications layer that allows applications to interact across hardware and network environments (ATG).
Migration	The process of moving or copying data files from one storage medium or storage entity to another; moving data from one system to another.
MIME type	Specifies how a file is created and how it is to be viewed; this information is contained in the filename extension, e.g., afile.doc or index.html, and in server MIME type files (JPL D-14206).
Model	A limited representation of a system or process; the role of a model is to answer questions about the entity it represents; model types may include executables, metadata, design, operations, process, enterprise, and organization (INCOSE); also a representation in software or hardware of one or more properties of a product and possibly its operational environment for the purpose of design, analysis, or verification—it is an idealized representation of a product (ISO 10303); see also *data model*.
Multi-dimensional database management system	Captures and presents data that can be arranged in multiple dimensions; users can analyze large amounts of data using an intuitive paradigm (ATG).
Object	An abstraction of a real-world thing; a typical but unspecified instance
Object-oriented database management system (ODBMS)	Information is organized as objects, classified by a class type and organized in a class family hierarchy.
Open architecture	An architecture based on open industry specifications (JPL D-12950).
Open system	A system using *open architecture.*
Operational system	A system that supports the daily, active operations of the enterprise or the group (ATG).
Product	Consists of one or more entities of various types; also a thing or substance produced by a natural or artificial process (ISO 10303).
Product data	A representation of information about a product in a formal manner suitable for communication, interpretation , or processing by a human being or a computer (ISO 10303).
Product information	Facts, concepts, or instructions about a product (ISO 10303).
Propagation	The process of keeping a copy of data usually for the purpose of synchronization during distribution of data from the source to the target database (ATG); a.k.a. replication.
Query	A request for information posed by the user or tool operated by the user (ATG); see also *ad hoc query*.
Record	A collection of information pertaining to an entity, usually divided into fields containing the values assigned to each attribute.

Term	Description
Reference designator	A unique numeric or alphanumeric identifier assigned to an entity.
Relational database management system (RDBMS)	A collection of data that are logically related; built on the relational model of tables, columns, and views.
Replication	The process of keeping a copy of data usually for the purpose of synchronization during distribution of data from the source to the target database (ATG); a.k.a. propagation.
Report	The result presented in response to a query or set of queries. Complex reports with customized formatting may need to be designed and built. Reports can be generated on-demand or scheduled to run at predefined intervals.
Repository	A place where information is stored; some online repositories may meet the criteria of an *information management system*.
Schema	A representation of the organization and relationships between data objects and their attributes. For example, a database schema identifies the table and column structures and relationships or dependencies between attributes that appear in more than one table.
Source database	The database from which data is to be extracted or copied (ATG).
Standard	A specification of a protocol, convention, procedure, or format that is developed and maintained by an independent standards body using an open development process (such as ISO, ANSI, IEEEE, or IAB). Sometimes referred to as a *de jure* standard, as opposed to a *de facto* standard which is a product or system that has captured a large share of the market and is emulated, copied, and used by others to obtain market share.
Target database	The database into which data will be loaded or inserted (ATG).
Transformation	The process of creating validated data by filtering, cleansing, merging, decoding, translating, denormalizing, converting, aggregating, and auditing (ATG).
Upload	The process of moving or copying a file or files from a workstation to a server
User	A consumer of information; also referred to as knowledge worker or end-user.
User interface	An interface of a system at which physical interactions take place between a human being and the system (ISO/IEC TR 10000-3); a.k.a. human-computer interface.
Web	A network of computer systems and subnets (part of the Internet) on which hypertext and other browsable formats are made available; short for World Wide Web (JPL D-12950).

Sources:

AIAI:	Artificial Intelligence Applications Institute
ATG:	Applied Technologies Group, Technology Guides, http://www.techguide.com
Dublin:	Dublin Core Metadata Element Set, http://purl.org/metadata/dublin_core
Internet Standards Process:	Rev. 3, RFC 2026
INCOSE:	International Council on Systems Engineering; formed to develop, nurture, and enhance the system engineering approach to multidisciplinary system product development.
ISO 10303:	Standard for the Exchange of Product Model Data (STEP)
ISO/IEC TR 10000-3:	Framework and Taxonomy of International Standardized Profiles—Principles and Taxonomy for OSE Profiles
JPL D-12950:	Engineering Data System Functional Requirements, February 1996
JPL D-14206:	Electronic Publishing Handbook, Enterprise Data Architecture Team
Telos:	Introduction to Telos Pangaea, Telos Information Systems Inc.